ON THE EQUIVALENCE OF OPTIMAL CONTROL

PROBLEMS AND THE TRANSFORMATION OF OPTIMAL

CONTROL PROBLEMS WITH COMPACT CONTROL

REGIONS INTO LAGRANGE PROBLEMS

By

STEPHEN KENT PARK

A thesis submitted to the Graduate Faculty of
North Carolina State University at Raleigh
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

DEPARTMENT OF MATHEMATICS

RALEIGH

1 9 7 0

APPROVED BY:

_____   _____

_____   _____
                                        Chairman

ABSTRACT

PARK, STEPHEN KENT. On the equivalence of optimal control problems and the transformation of optimal control problems with compact control regions into Lagrange problems. (Under the direction of HANS SAGAN.)

Several natural types of equivalence are defined for a set of optimization problems. Using these concepts of equivalence, a method is described by which an optimal control problem with a compact control region U can be transformed into an equivalent optimal control problem with an (in general) arbitrary control region Z. This method assumes the existence of a function ψ which maps Z onto U and satisfies certain continuity requirements. It is shown that when U is convex the existence of this function is guaranteed and that when U is, in addition, a polyhedron various representations for ψ are given.

The particular case where Z is all of euclidean p-space, for some positive integer p, is considered in some detail. It is shown that when U is a convex body the optimal control problem is equivalent to (i.e., may be transformed into) an optimal control problem in which the control functions are unconstrained. Furthermore, when U is a polyhedron it is shown that the optimal control problem is equivalent to a classical Lagrange problem. This later result provides, as an application, a proof of the Pontryagin maximum principle based upon known necessary conditions in the calculus of variations.

BIOGRAPHY

Stephen Kent Park was born in Washington D.C. on October 12, 1942. He was educated in the public school system of Pennsylvania and graduated from Morrison Cove High School, Martinsburg, Pennsylvania in 1960.

In 1964 he received the Bachelor of Science degree in Mathematics from Shippensburg State College, Shippensburg, Pennsylvania and began work for the National Aeronautics and Space Administration at Langley Research Center, Hampton, Virginia. As an employee of the NASA he commenced full time graduate study at North Carolina State University in 1966 and received a Master of Applied Mathematics in 1967. In 1969, upon the completion of graduate study, he returned to the Langley Research Center where he is presently employed in the Analysis and Computation Division.

In August of 1968 he married Miss Rosa V. Hopkins of Raleigh, North Carolina, and he and his wife presently reside in Newport News, Virginia.

ACKNOWLEDGMENTS

The author wishes to express his thanks to the National Aeronautics and Space Administration for its cooperation and financial support throughout his post graduate study and research. In particular, thanks are extended to his former supervisor, George P. Wood, for continued cooperation and to his present supervisor, John E. Hogge, for his encouragement and understanding. To the typist, Mrs. Judith T. Burton, the author offers not only his thanks but also his sympathy - may she never again have to type another $\alpha:[t_0, t_1] \to R^\mu$, $\Omega_2^O \subset \Omega_2^{OO} \subset \Omega_2$ or $\psi(B^m(1, 0)) = [-1, 1]^m$.

Special gratitude is extended to North Carolina State University and to the members of the author's Advisory Committee: Doctors R. A. Struble, K. Koh, and G. H. Katzin. To the committee chairman, Doctor Hans Sagan, the author expresses sincere appreciation for the continued interest and support which have helped to make the past four years of study and research a rewarding and enjoyable experience.

To my wife, Rosa, who by her own admission endured no great hardships or undue burdens during this period, the author says thank you for the help rendered by successfully performing the hardest job of all - being a good wife.

TABLE OF CONTENTS

LIST OF FIGURES

1. INTRODUCTION

On page 218 of (22), L. C. Young expresses the following sentiment.
"We see that there is no real difference between the Lagrange problem and
optimal control; the latter is simply a more up-to-date formulation.
Apparent minor differences are sometimes pointed out, but they are really
quite insignificant." It is interesting to contrast this statement with the
one made by L. S. Pontryagin et al. on page 239 of (18) which reads
". . . the optimal control problem is a generalization of the problem of
Lagrange in the calculus of variations, and is equivalent to the latter in
the case where the control region U is an open set . . ." Regardless
of which of these two philosophies one chooses to adopt, when faced with an
optimal control problem with a (typically) closed and bounded control region
it is not at all clear, in most cases, what results from the calculus of
variations (if any) are applicable to the problem. As early as 1948,
M. R. Hestenes (8) used the calculus of variations to analyze an optimal
control system in which the control region was defined by a system of
inequalities. He used the method of "slack variables" - usually attributed
to F. A. Valentine (21) - and transformed the optimal control problem into a
Bolza problem. Since that time, this approach has been investigated by a
variety of authors and a large portion of the research is contained in
Hestenes' book (9).

Herein we adopt a procedure for transforming optimal control problems
into calculus of variation problems which is similar in philosophy to
Hestenes', but which makes no use of slack variables. The procedure is

highly geometric in flavor and places an emphasis on the construction of mappings from one control region onto another. Using these mappings, we define a rather natural type of equivalence in the set of all optimal control problems (of a certain type). As an immediate by-product we find that it is possible to identify a large class of optimal control problems which are, to within a transformation, Lagrange problems in the calculus of variations.

The procedure mentioned in the previous paragraph may be illustrated as follows. Consider the problem of optimally steering a control system with the control function u subject to the constraint $u(t) \in U$. Intuitively, one would believe that, at least if ψ is continuous, one-to-one, onto and has a continuous inverse, this problem is equivalent to the problem of optimally steering the same system with a control function z subject to the constraint $z(t) \in Z$ where $u(t) = \psi(z(t))$ and ψ maps Z onto U. This idea is very briefly mentioned on page 218 of (22) and Young picturesquely concludes that "this corresponds to providing a new set of dials z, which are used to control the original dials u, as in regulating a television set without getting up from one's arm chair."

To be more specific, the basic problem with which we will be concerned is the following one. Suppose an optimal control system is given in which the control (vector) function is constrained to lie in a compact control region contained in euclidean m-space R^m. We seek to determine conditions sufficient to guarantee that there exists a positive integer p, a subset Z of R^p, and a function ψ from Z onto U which together transform the original control system into a new control system, with control region Z,

which is (in some sense) equivalent to the original. Having done this we seek additional conditions which will permit us to choose $Z = R^p$ and thereby transform the original optimal control problem into an equivalent Lagrange problem.

In chapter 2 we consider a very broad class of optimization problems and discuss two concepts of equivalence in this class. In anticipation of investigating the questions mentioned in the previous paragraph, we also define what is called, for lack of a better name, Ψ-equivalence. It turns out that Ψ-equivalence is stronger than one of the previously mentioned concepts of equivalence; namely, if an optimization problem P_2 is Ψ-equivalent to an optimization problem P_1 then P_2 and P_1 are equivalent. In the context of the previous paragraph Ψ-equivalence is a very natural concept. For it turns out that if the original optimal control problem, say P_1, is transformed into a second problem P_2 and if P_2 is Ψ-equivalent to P_1 then the following is true: if $u:[t_0, t_1] \to U$ is a (bounded, measurable) optimal control function for P_1 then there exists at least one (bounded, measurable) optimal control function $z:[t_0, t_1] \to Z$ for P_2 such that $u = \Psi \circ z$ (where $\circ$ denotes function composition); conversely if $z:[t_0, t_1] \to Z$ is an optimal control for P_2 then $u = \Psi \circ z$ is an optimal control for P_1. Therefore, we ask the question, under what conditions is the transformed problem P_2 Ψ-equivalent to the original problem P_1. Theorem 2.2 and corollary 2.1 give a complete answer to this problem i.e., necessary and sufficient conditions for P_2 to be Ψ-equivalent to P_1. From these results we obtain corollary 2.2 which serves as the basis of the investigation in chapter 3.

The principle result of chapter 3 is theorem 3.4 which relies heavily on Filippov's implicit function lemma (5) and corollary 2.2. The theorem says that if there exists a positive integer p, a set $Z \subset R^p$, and continuous function $\psi : Z \to U$ which satisfies $\psi(Z) = \psi(Z^*) = U$ where $Z^* \subset Z$ is compact, then P_2 is Ψ-equivalent to P_1. It is a tribute to the strength of Filippov's lemma that theorem 3.4 remains true without any reference to the structure of the sets $\psi^{-1}(u)$ where $u \in U$ (except, of course, for the structure of $\psi^{-1}(u)$ induced by the continuity of ψ). Specifically ψ is not by necessity one-to-one which illustrates that, in the space of control functions, the normal (i.e., ψ is a homeomorphism) concept of a change in variables is definitely overly restrictive.

Chapter 4 is an application of theorem 3.4 with $Z = R^p$ and ψ continuously differentiable. Specifically, we show that in many cases there exists a Lagrange problem which is Ψ-equivalent to the original optimal control problem. In chapter 5 we use the results of chapter 4 to explore further the relationships between the theory of optimal control and the calculus of variations. The principle result here is that, using the multiplier rule, Weierstrass condition, transversality conditions and the concept of Ψ-equivalence, we are able to obtain, for certain compact control regions U, a proof of the Pontryagin maximum principle. Since this proof serves as an illustration of the material contained in chapters 2, 3, 4, and 6, it is by necessity not ideally organized. Certainly if our sole objective had been a proof of the maximum principle much of the material in each chapter could have been eliminated and the remaining material reorganized.

Since the material of chapters 2 through 5 rests on the assumption that a set Z and an onto function $\psi : Z \to U$ exist, we turn to this question in chapter 6. We show that when U is a convex body in R^m then the hypotheses of theorem 3.4 are satisfied. In the less general case where U is a polyhedron we explicitly construct such a ψ which is, in addition, continuously differentiable with $Z = R^p$. Finally in the simple case where U is a parallelipiped we show that ψ assumes a very simple form.

In the appendix we consider the question of local, as opposed to global, ψ-equivalence. The discussion parallels chapter 2 and we obtain the local analogues of theorem 2.2 and corollary 2.1. We close with a brief discussion of an application of the transformation approach outlined herein to the problem of minimizing a real valued function of n variables.

Since the idea of a change of variables in the space of control functions is such an intuitive one, it is reasonable to assume that it has been considered by others. This is, in fact, true for, as pointed out in chapter 6, it is common practice (for example) to begin the statement of an optimal control problem with a phrase such as "let the control region U be a right parallelipiped and assume without loss of generality that $U = [-1, 1]^m$." This statement is based on the (implicit) fact that there is a (linear) homeomorphism from $[-1, 1]^m$ onto the right parallelipiped and the (implicit) assumption that a change of variables induced by a homeomorphism preserves equivalence. The idea of transforming an optimal control problem into a Lagrange problem (in the manner outlined herein) is far less frequently encountered. However, the idea is mentioned, for example, in (2) and (14) and on pages 134-138 of (4). Typically these references use $U = [-1, 1]$

with $Z = R$ and $\psi(z) = \sin z$ (or the immediate generalization of this idea to the case $U = [-1, 1]^m$) coupled with the (implicit) assumption that the transformed problem is equivalent to the original.

2. THE EQUIVALENCE OF OPTIMIZATION PROBLEMS

Let Λ be a suitable index set and suppose that for each $\lambda \in \Lambda$ there corresponds a set Y_λ, a nonempty subset $\Omega_\lambda \subset Y_\lambda$, and a functional $F_\lambda : \Omega_\lambda \to R$ where R denotes the reals. Despite the fact that Ω_λ may have no algebraic or topological structure one can still consider the following global problem: to find an $\alpha^O \in \Omega_\lambda$ such that $F_\lambda(\alpha^O) \leq F_\lambda(\alpha)$ for all $\alpha \in \Omega_\lambda$. Henceforth, for $\lambda \in \Lambda$, we will refer to this problem as P_λ, refer to the elements of Ω_λ as admissible elements and refer to the element $\alpha^O \in \Omega_\lambda$ (if it exists) as a global optimal (admissible) element. Furthermore, if a global optimal element (elements) exists for P_λ, we will frequently refer to it (them) as a solution (solutions) although in the case where no global optimal elements exist the term solution denotes, instead, a proof of this fact. Finally, since we will always be concerned with global (as opposed to local) properties we adopt the convention that in the absence of a modifier such terms as optimal element, optimization problem and equivalence should be understood in the global sense.[1]

The formulation of the preceding problem is sufficiently general to guarantee that a wide variety of what could be loosely referred to as optimization or minimization problems are included in the set $\Gamma = \left\{ P_\lambda : \lambda \in \Lambda \right\}$. Specifically Γ contains the rather general optimal control problem to be formulated in chapter 3 and the Lagrange problem to be formulated in chapter 4. When discussing optimal control and/or Lagrange problems the notion of

---

[1] In the appendix the material of this chapter is interpreted in terms of a local problem.

equivalence of problems P_λ and P_μ for $\lambda, \mu \in \Lambda$ is frequently mentioned - in several different contexts. Unfortunately, these notions of equivalence are often only vaguely (if at all) defined. Furthermore, equivalence may not mean equivalence in the strict mathematical sense, i.e., in the sense that P_λ and P_μ are equivalent, written $P_\lambda \equiv P_\mu$, if $\equiv$ is an equivalence relation on Γ. With this in mind let us first define and then discuss the following equivalence concepts.

Definition 2.1: For $\lambda \in \Lambda$, let $\Omega_\lambda^O \subset \Omega_\lambda$ denote the set of optimal elements for problem P_λ.

Definition 2.2: For $\lambda, \mu \in \Lambda$ if there exists a function $\Psi : \Omega_\mu \to \Omega_\lambda$ such that $F_\mu = F_\lambda \circ \Psi$ then

(1) P_μ is strictly equivalent to P_λ if Ψ is one-to-one and
$\Psi(\Omega_\mu) = \Omega_\lambda$

(2) P_μ Ψ-equivalent to P_λ if $\Psi(\Omega_\mu) = \Omega_\lambda$

(3) P_μ is weakly Ψ-equivalent to P_λ if $\Psi(\Omega_\mu^O) = \Omega_\lambda^O$.

In other words P_μ is strictly equivalent to P_λ if Ψ is a one-to-one correspondence between the admissible elements of P_μ and P_λ, P_μ is Ψ-equivalent to P_λ if Ψ maps the admissible elements for P_μ onto the admissible elements for P_λ, P_μ is weakly Ψ-equivalent to P_λ if Ψ maps the optimal elements for P_μ onto the optimal elements for P_λ, and in each of these cases, corresponding admissible elements have the same functional value. As an immediate consequence of this definition we have the following.

Remark 2.1: Strict equivalence is an equivalence relation in Γ. Thus we have that the ordering of P_λ and P_μ in the statement " P_μ is

strictly equivalent to P_λ " is inconsequential. In contrast to this situation we see that Ψ-equivalence and weak Ψ-equivalence do not define equivalence relations on Γ. However, it is true that for $\lambda \in \Lambda$, P_λ is both Ψ-equivalent and weakly Ψ-equivalent to itself (Ψ is the identity on Ω_λ). Also the following is true.

Remark 2.2: For $\lambda, \mu, \nu \in \Lambda$ if P_μ is Ψ-equivalent (weakly) to P_λ and P_ν is Φ-equivalent (weakly) to P_μ then P_ν is $\Psi \circ \Phi$-equivalent (weak) to P_λ where $\circ$ denotes function composition.

In other words the concepts of Ψ-equivalent and weak Ψ-equivalence define relations in Γ which are reflexive and transitive but are not, in general, symmetric. It is for this reason that throughout we will make a careful distinction between the statement "P_λ and P_μ are strictly equivalent" and the weaker statements "P_μ is Ψ-equivalent (weakly Ψ-equivalent) to P_λ" - in addition we will be careful to observe the order of terms P_μ and P_λ in the later statements.

It is clear that strict equivalence is a stronger concept than Ψ-equivalence. The following theorem says that conversely if P_μ is Ψ-equivalent to P_λ then there exists a "subproblem" P_ν contained in P_μ which is strictly equivalent to P_λ.

Theorem 2.1: If P_μ is Ψ-equivalent to P_λ then there exists a nonempty subset $\Omega_\nu \subset \Omega_\mu$ such that Ω_ν and $F_\nu = F_\mu \big| \Omega_\nu$ (i.e., the restriction of F_μ to Ω_ν) define a problem P_ν which is strictly equivalent to P_λ.

Proof: Since $\Psi:\Omega_\mu \to \Omega_\lambda$ is onto, for each $\alpha \in \Omega_\lambda$, the set $\Psi^{-1}(\alpha)$ is nonempty. Hence (by the axiom-of-choice) there exists a function $\Phi:\Omega_\lambda \to \cup\{\Psi^{-1}(\alpha):\alpha \in \Omega_\lambda\}$ such that $\Phi(\alpha)\in \Psi^{-1}(\alpha)$ for each $\alpha \in \Omega_\lambda$ i.e., $\Psi(\Phi(\alpha)) = \alpha$. Let $\Omega_\nu = \{\Phi(\alpha):\alpha \in \Omega_\lambda\}$ then by construction $\Omega_\nu \subset \Omega_\mu$ and $\Omega_\nu \neq \emptyset$. Furthermore, the function Φ is one-to-one since $\Phi(\alpha) = \Phi(\alpha^*)$ implies $\alpha = \Psi(\Phi(\alpha)) = \Psi(\Phi(\alpha^*)) = \alpha^*$. Let $F_\nu = F_\mu\big|_{\Omega_\nu}$ and thereby define P_ν. If $\alpha \in \Omega_\lambda$ then $\Phi(\alpha)\in \Omega_\nu \subset \Omega_\mu$ and hence

$$F_\nu(\Phi(\alpha)) = F_\mu(\Phi(\alpha)) = F_\lambda(\Psi(\Phi(\alpha)) = F_\lambda(\alpha)$$

i.e., $F_\nu \circ \Phi = F_\lambda$. Since $\Phi(\Omega_\lambda) = \Omega_\nu$, the theorem is proved.

Let us turn to the concept of weak Ψ-equivalence. Following some preliminary lemmas we obtain (theorem 2.2) an alternate characterization of weak Ψ-equivalence and as a corollary we obtain the expected result that Ψ-equivalence is indeed stronger than weak Ψ-equivalence. Hereafter let us avoid reference to the general index set Λ and instead simply refer to the optimization problems as P_1, P_2 and so forth. Furthermore, in the hypotheses of the following three lemmas let P_2 be defined by Ω_2 and $F_2 = F_1 \circ \Psi$ where $\Psi:\Omega_2 \to \Omega_1$ (Ψ is not necessarily onto).

Lemma 2.1: If $\alpha \in \Omega_1$, $\beta \in \Omega_2$ and $\Psi(\beta) = \alpha$ then $\alpha \in \Omega_1^o$ implies $\beta \in \Omega_2^o$.

Proof: Consider $\beta^* \in \Omega_2$ and $\alpha^* = \Psi(\beta^*)\in \Omega_1$. By hypothesis $F_1(\alpha^*) \geq F_1(\alpha)$ and thus

$$F_2(\beta^*) = F_1(\Psi(\beta^*)) = F_1(\alpha^*) \geq F_1(\alpha) = F_1(\Psi(\beta)) = F_2(\beta)$$

which establishes the lemma.

Lemma 2.2: $\Psi(\Omega_2^O) \supset \Omega_1^O$ iff $\Psi(\Omega_2) \supset \Omega_1^O$.

Proof: If $\Omega_1^O = \emptyset$ the lemma is trivially true. Therefore assume $\Omega_1^O \neq \emptyset$ and consider $\alpha^O \in \Omega_1^O$. If $\Psi(\Omega_2) \supset \Omega_1^O$ then there exists $\beta^O \in \Omega_2$ such that $\Psi(\beta^O) = \alpha^O$. From lemma 2.1 we have $\beta^O \in \Omega_2^O$ and thus $\Psi(\Omega_2^O) \supset \Omega_1^O$. Conversely $\Psi(\Omega_2^O) \supset \Omega_1^O$ implies $\Psi(\Omega_2) \supset \Omega_1^O$ immediately since $\Omega_2 \supset \Omega_2^O$.

Lemma 2.3: $\Omega_2^O \neq \emptyset$ and $\Psi(\Omega_2^O) \subset \Omega_1^O$ iff $\Omega_1^O \neq \emptyset$ and $\Psi(\Omega_2) \cap \Omega_1^O \neq \emptyset$.

Proof: If $\Omega_1^O \neq \emptyset$ and $\Psi(\Omega_2) \cap \Omega_1^O \neq \emptyset$ then there exists an $\alpha^O \in \Omega_1^O$ and $\beta^O \in \Omega_2$ such that $\Psi(\beta^O) = \alpha^O$. From lemma 2.1 $\beta^O \in \Omega_2^O$ i.e., $\Omega_2^O \neq \emptyset$. Furthermore, if $\beta \in \Omega_2^O$ then $F_2(\beta) = F_2(\beta^O)$. Letting $\alpha = \Psi(\beta)$ then $F_1(\alpha) = F_2(\beta) = F_2(\beta^O) = F_1(\alpha^O)$. Since $\alpha^O \in \Omega_1^O$ this implies $\alpha \in \Omega_1^O$ and thus $\Psi(\Omega_2^O) \subset \Omega_1^O$. Conversely, if $\Omega_2^O \neq \emptyset$ and $\Psi(\Omega_2^O) \subset \Omega_1^O$ then there exists $\beta^O \in \Omega_2^O$ and thus $\alpha^O = \Psi(\beta^O) \in \Omega_1^O$. Therefore, $\Psi(\Omega_2^O) \cap \Omega_1^O \neq \emptyset$ which implies $\Psi(\Omega_2) \cap \Omega_1^O \neq \emptyset$.

Combining lemmas 2.2 and 2.3 we obtain the following theorem which gives a necessary and sufficient condition for P_2 to be weakly Ψ-equivalent to P_1.

Theorem 2.2: Suppose $\Omega_1^O \neq \emptyset$. P_2 is weakly Ψ-equivalent to P_1 iff $\Psi(\Omega_2) \supset \Omega_1^O$. If instead $\Omega_1^O = \emptyset$ then P_2 is weakly Ψ-equivalent to P_1 iff $\Omega_2^O = \emptyset$.

Proof: If $\Omega_1^O = \emptyset$ the theorem follows as a restatement of definition 2.2. If $\Omega_1^O \neq \emptyset$ and P_2 is weakly Ψ-equivalent to P_1 then $\Psi(\Omega_2^O) = \Omega_1^O$ i.e., $\Psi(\Omega_2) \supset \Omega_1^O$ since $\Omega_2 \supset \Omega_2^O$. Conversely, if $\Omega_1^O \neq \emptyset$ and $\Psi(\Omega_2) \supset \Omega_1^O$ then from lemma 2.2 $\Psi(\Omega_2^O) \supset \Omega_1^O$. Furthermore, $\Omega_1^O \neq \emptyset$ implies $\Psi(\Omega_2) \cap \Omega_1^O \neq \emptyset$ and thus from lemma 2.3 $\Psi(\Omega_2^O) \subset \Omega_1^O$ i.e., P_2 is weakly Ψ-equivalent to P_1.

Corollary 2.1: If P_2 is Ψ-equivalent to P_1 then P_2 is weakly Ψ-equivalent to P_1.

Proof: If $\Omega_1^O \neq \emptyset$ then theorem 2.2 applies since $\Omega_1^O \subset \Omega_1$. If $\Omega_1 = \emptyset$ suppose there exists $\beta^O \in \Omega_2^O$ and let $\alpha^O = \Psi(\beta^O)$. Consider $\alpha \in \Omega_1$ then by hypothesis there exists $\beta \in \Omega_2$ such that $\Psi(\beta) = \alpha$. Since $F_2(\beta) \geq F_2(\beta^O)$ this implies $F_1(\alpha) \geq F_1(\alpha^O)$ and since α was arbitrary it follows that $\alpha^O \in \Omega_1^O$. This contradiction establishes $\Omega_2^O = \emptyset$ and from theorem 2.2, P_2 is weakly Ψ-equivalent to P_1.

As mentioned in the introduction our primary interest herein will be concerned (in general) with transforming a given optimal control problem (P_1) into a second problem (P_2) which is in some sense equivalent to the former. Before considering this question in detail let us consider one final concept of equivalence.

Definition 2.3: Problems P_1 and P_2 will be said to be equivalent if there exist functions $\Psi_{12}:\Omega_1 \to \Omega_2$ and $\Psi_{21}:\Omega_2 \to \Omega_1$ which satisfy $F_2 = F_1 \circ \Psi_{21}$ and $F_1 = F_2 \circ \Psi_{12}$.

Notice that this does, in fact, define an equivalence relation on Γ. Furthermore, if P_1 and P_2 are strictly equivalent (i.e., there exists a one-to-one, onto function $\Psi:\Omega_2 \to \Omega_1$ satisfying $F_2 = F_1 \circ \Psi$) then using $\Psi_{21} = \Psi$ and $\Psi_{12} = \Psi^{-1}$ we have $F_1 = F_2 \circ \Psi^{-1}$ and P_1, P_2 are equivalent. That is, strict equivalence is stronger than equivalence. As we observed previously, Ψ-equivalence is not an equivalence relation on Γ but in the sense of theorem 2.1, Ψ-equivalence is "almost" strict equivalence. From the proof of theorem 2.1 we obtain the following result.

Theorem 2.3: If P_2 is Ψ-equivalent to P_1 then P_1 and P_2 are equivalent (in the sense of definition 2.3).

Proof: Referring to the proof of theorem 2.1 (with $1 = \lambda$, $2 = \mu$, $3 = \nu$) we see that $\Psi : \Omega_2 \to \Omega_1$ and $\Phi : \Omega_1 \to \Omega_3 \subset \Omega_2$ satisfy the conditions of definition 2.3.

With regard to the previously mentioned problem of, given P_1, constructing a problem P_2 which is in some sense equivalent to P_1 the procedure which we will pursue is clear. Starting with P_1, we construct a set Ω_2 and a function $\Psi : \Omega_2 \to \Omega_1$. If we define $F_2 : \Omega_2 \to R$ as $F_2 = F_1 \circ \Psi$ and if the function Ψ is onto, i.e., $\Psi(\Omega_2) = \Omega_1$ then P_2 is Ψ-equivalent to P_1 and hence P_1, P_2 are equivalent. Of more importance is the fact that P_2 is weakly Ψ-equivalent to P_1 and as a consequence the following is true: if $\Omega_1^O = \emptyset$ then $\Omega_2^O = \emptyset$; if $\Omega_1^O \neq \emptyset$ then for each $\alpha^O \epsilon \, \Omega_1^O$ there exists at least one $\beta^O \, \epsilon \, \Omega_2^O$ such that $\Psi(\beta^O) = \alpha^O$; finally if $\beta^O \, \epsilon \, \Omega_2^O$ then $\Psi(\beta^O) \epsilon \, \Omega_1^O$ i.e., $\Psi(\beta^O)$ is an optimal element for P_1. Thus, if we analyze P_2 we have, to within a transformation (namely $\Psi : \Omega_2 \to \Omega_1$), at the same time analyzed P_1. Since we will refer to this idea repeatedly, we record the previous discussion as the FUNDAMENTAL CONSTRUCTION LEMMA which is the following.

Corollary 2.2: If there exists a set Y_2, a nonempty set $\Omega_2 \subset Y_2$ and a function $\Psi : \Omega_2 \to \Omega_1$ with Ψ onto (i.e., $\Psi(\Omega_2) = \Omega_1$) then (Y_2, Ω_2, F_2) define a problem P_2 which is Ψ-equivalent to P_1 where $F_2 = F_1 \circ \Psi$.

It is this corollary that will be the basis of the discussion in chapters 3 and 4.

3. Ψ-EQUIVALENCE IN A CLASS OF OPTIMAL CONTROL PROBLEMS

The discussion in this chapter will be based upon the abstract theory of Ψ-equivalence developed in chapter 2. Specifically, we will be concerned with corollary 2.2 and its application to the theory of optimal control. To avoid lengthy introductory remarks we adopt the convention that, unless explicitly stated otherwise, the terminology and notational conventions of Pontryagin et al. (18) will be used throughout. (Although with regard to vector notation, we follow a convention opposite to theirs concerning subscripts and superscripts.)

3.1 The Formulation of Problem P_1

For any positive integer r let R^r denote euclidean r-space and let n and m be fixed positive integers. Suppose that the phase space X is all of R^n and that an initial point a and nonempty target set S are given in X along with an initial time $t_0 \in R$. In addition let the control region U be a nonempty subset of R^m. We assume that S is closed, $a \notin S$, and U is compact (i.e., closed and bounded). Finally, we assume the real valued functions g_i and $\dfrac{\partial g_i}{\partial x_j}$ are defined and continuous on $X \times U$ for $i = 0, 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$ and we let $g = (g_1, \ldots, g_n)$. The optimal control problem designated P_1 will be formulated in terms of the quantities mentioned in this paragraph. In anticipation of this, define the following:

Definition 2.1: $Y_1 = \left\{ u = (u_1, \ldots, u_m) : u \text{ is measurable and} \right.$

$$u : [t_0, t_*] \to U \text{ for some } t_* > t_0 \Big\}$$

We remark that measurable and integrable are meant in the Lebesgue sense and that a vector function is measurable (integrable) if each of its components is measurable (integrable). Furthermore, since U is bounded it follows that if $u \in Y_1$ then u is integrable.

The following lemma is a standard result in analysis (see, e.g., (19) page 39), but since it will be used repeatedly we choose to label it lemma 3.1 to facilitate later reference to its contents.

Lemma 3.1: If $\alpha: A \to R^\mu$ is continuous for $\emptyset \neq A \subset R^\nu$ and positive integers ν, μ then for each bounded, measurable function $\beta: [t_0, t_1] \to A$ the function $\alpha \circ \beta: [t_0, t_1] \to R^\mu$ is integrable.

Returning to the optimal control problem let us define some standard terminology.

Definition 3.2: If corresponding to a finite $t_1 > t_0$ and function $u: [t_0, t_1] \to U$ in Y_1 there exists an absolutely continuous function $x: [t_0, t_1] \to X$ such that $\dot{x}(t) = g(x(t), u(t))$ a.e. on $[t_0, t_1]$, $x(t_0) = a$ and $x(t_1) \in S$ then u is an admissible control (function), x is an admissible trajectory, and t_1 is a terminal time.

Since it requires $x(t_1) \in S$, this definition of an admissible control is stronger than the one given by Pontryagin et al. However (as observed by many authors), a control which does not ultimately steer the trajectory to the target set cannot be an optimal control and thus there seems to be little value in considering a control which is not admissible in the sense of definition 3.2 for some $t_1 > t_0$. With regard to definition 3.2 the next lemma says that the admissible trajectory associated with an admissible

control is unique and that without loss of generality we may take t_1 as

the least terminal time. The uniqueness of the trajectory is a standard

result in the theory of differential equations (see (18), p. 78) and thus

we only concern ourselves with a proof that there exists a $t_1 > t_0$ such

that $x(t_1) \in S$ and $t \in [t_0, t_1)$ implies $x(t) \notin S$.

    <u>Lemma 3.2</u>: If $u:[t_0, t_*] \to U$ is an admissible control and $x:[t_0, t_*] \to X$

is an associated admissible trajectory then $t_1 = \inf \left\{ \tau \geq t_0 : x(\tau) \in S \right\} > t_0$

is a terminal time and x restricted to $[t_0, t_1]$ is unique.

    <u>Proof</u>: Let $T = \left\{ \tau \geq t_0 : x(\tau) \in S \right\}$ then $t_* \in T$ and thus

$t_1 = \inf T \geq t_0$ exists. By definition there exists a sequence $(\tau_n) \to t_1$

with $\tau_n \in T$, <u>i.e.</u>, $x(\tau_n) \in S$ and $\tau_n \geq t_0$. Furthermore, x is continuous

on $[t_0, t_*]$ and in particular is continuous at t_1 since $t_0 \leq t_1 \leq t_*$.

Therefore $(x(\tau_n)) \to x(t_1)$ and since S is closed it follows that

$x(t_1) \in S$ <u>i.e.</u>, t_1 is a (least) terminal time. Finally, if $t_0 = t_1$ then

$x(t_1) = x(t_0) = a$ which contradicts the assumption $a \notin S$. Thus $t_0 < t_1$.

    <u>Lemma 3.3</u>: If u is an admissible control with the corresponding

(unique) admissible trajectory x and (least) terminal time t_1 then

$$F_1(u) = \int_{t_0}^{t_1} g_0(x(\tau), u(\tau)) \, d\tau$$

exists.

    Proof: Since x and u are bounded, measurable on $[t_0, t_1]$ and since g_0 is

continuous on $X \times U$, the existence of $F_1(u)$ follows from lemma 3.1.

In other words, lemma 3.3 serves to define a functional F_1 on the set of admissible control functions. In agreement with the notation of chapter 2 we make the following definition:

Definition 3.3: Let $\Omega_1 \subset Y_1$ denote the <u>set of admissible control functions</u>. We will sometimes write, for brevity, $(u; x, t_1) \in \Omega_1$ if $u \in \Omega_1$, x is the associated trajectory, and t_1 is the terminal time.

In passing we note that it may seem more consistent with the notation of chapter 2 to define Ω_1 as $\{\alpha : \alpha = (u, x, t_1)\}$ where u is an admissible trajectory, x is the associated trajectory, and t_1 is the terminal time. However, since x and t_1 are unique (for each u) this is unnecessary and in fact is undesirable since it would require us to adopt some clumsy conventions concerning the construction of Y_1. Therefore we use the definition of Ω_1 given by definition 3.3 and yield to the previous discussion by occasionally writing $(u; x, t_1) \in \Omega_1$.

The usual purpose of optimal control theory is not to determine whether or not Ω_1 is empty, but rather, given $\Omega_1 \neq \emptyset$, to find a minimal element of Ω_1. Therefore we <u>assume</u> $\Omega_1 \neq \emptyset$ and state the following.

Definition 3.4: Problem P_1 is: find a $u^o \in \Omega_1$ such that $F_1(u^o) \leq F_1(u)$ for all $u \in \Omega_1$. If an optimal element $(u^o; x^o, t^o_1)$ exists for P_1 we call u^o an <u>optimal control</u> (function) and x^o the associated <u>optimal trajectory</u>. Finally, we denote the set of optimal control functions as Ω^o_1.

3.2 The Construction of Problem P_2

Consistent with the discussion in chapter 2 we have defined P_1 and now we make the following basic assumption which we assume is satisfied throughout the remainder of this chapter. A later chapter (chapter 6) is devoted to the question of the existence and explicit construction of a set Z and function ψ satisfying the basic assumption for particular choices of the set U.

Basic assumption: there exists the following

(1) a positive integer p,

(2) a nonempty set $Z \subset R^p$ (perhaps $Z = R^p$),

(3) and a <u>continuous</u> function $\psi: Z \to U$ with $\psi(Z) = U$, <u>i.e.</u>, ψ is onto.

It will be this set Z and function ψ which will serve as a basis for the construction of a set Ω_2, a function $\Psi: \Omega_2 \to \Omega_1$ and problem P_2 which is Ψ-equivalent to P_1.

<u>Definition 3.5</u>: For $(x, z) \in X \times Z$ define f_0 and $f = (f_1, \ldots, f_n)$ as

$$f_i(x, z) = g_i(x, \psi(z))$$

where $i = 0, 1, 2, \ldots, n$.

The following lemma is an immediate consequence of the properties of g_i and the continuity of ψ.

<u>Lemma 3.4</u>: The real functions f_i and $\dfrac{\partial f_i}{\partial x_j}$ are defined and continuous on $X \times Z$ for $i = 0, 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$.

The functions f_0 and $f = (f_1, \ldots, f_n)$ satisfy the same hypotheses as g_0 and g. Thus in terms of (a, t_0), X, S and Z we can obtain the

following results and definitions by formally replacing g_i by f_i and U by Z in the previous discussion. It is important to note that we have not assumed Z to be bounded and as a consequence some changes must be made.

Remark 3.1: $Y_2 = \{ z = (z_1, \ldots, z_p) : z$ is _bounded_, measurable and $z:[t_0, t_*] \to Z$ for some $t_* > t_0 \}$.

Remark 3.2: An admissible control $z:[t_0, t_1] \to Z$, admissible trajectory $x:[t_0, t_1] \to X$ and terminal time t_1 are defined analogous to definition 2.2 with the understanding that z is bounded with $\dot{x}(t) = f(x(t), z(t))$ a.e. on $[t_0, t_1]$.

With this understanding that z be bounded, the analogue of lemma 3.2 remains true (see (18), p. 78) so that corresponding to an admissible control z, the associated trajectory x is unique and again we take t_1 as the least terminal time. Furthermore, we sometimes write $(z; x, t_1) \in \Omega_2$ where:

Remark 3.3: Ω_2 denotes the set of (bounded) admissible control functions.

As the analogue of lemma 3.3 we have the following lemma which defines the functional F_2 on Ω_2.

Lemma 3.5: If $(z; x, t_1) \in \Omega_2$ then

$$F_2(z) = \int_{t_0}^{t_1} f_0 (x(\tau), z(\tau)) \, d\tau$$

exists.

Proof: From lemma 3.4 f_0 is continuous on $X \times Z$ and again since x and z are bounded, measurable lemma 3.5 follows from lemma 3.1.

While it may not be clear that $\Omega_2 \neq \emptyset$ we will show later that under the assumptions made so far this is indeed true. With this in mind, we make the following remark.

Remark 3.4: Problem P_2 is: find a $z^O \in \Omega_2$ such that $F_2(z^O) \leq F_2(z)$ for all $z \in \Omega_2$. If $(z^O;\, x^O,\, t_1^O)$ is an optimal element for P_2 then z^O and x^O are called an optimal control and optimal trajectory, respectively. The set of optimal control functions is designated Ω_2^O.

Following the program of chapter 2 we will show that P_2 as constructed is Ψ-equivalent to P_1 for an appropriate function $\Psi: \Omega_2 \to \Omega_1$. As one would expect, the function Ψ, which maps a function z to a function u is determined by applying the function ψ at each point $z(t)$. This idea is explicitly defined (definition 3.7) later but first some preliminary work is necessary.

Remark 3.5: If $(z;\, x,\, t_1) \in \Omega_2$ and the bounded function $z^*: [t_0,\, t_1] \to Z$ satisfies $z = z^*$ a.e. then $(z^*;\, x,\, t_1) \in \Omega_2$ and $F_2(z) = F_2(z^*)$. An analogous statement is true in Ω_1.

Remark 3.5 is an immediate consequence of the definition of an admissible trajectory and terminal time. It is merely an expression of the intuitive idea that changing a control function on a set of measure zero produces no effect on the trajectory and functional value. For this reason one can avoid distinguishing between two admissible controls which agree except on a set of measure zero.

The following two theorems express the fact that if the admissible controls z (for P_2) and u (for P_1) are related by $\psi \circ z = u$ then they produce the same admissible trajectory, terminal time, and functional value (i.e., $F_2(z) = F_1(u)$). Thus, in a generalized sense, P_2 is P_1 with a change of variables in the space of admissible controls. We will return to a discussion of this statement following theorem 3.3.

<u>Theorem 3.1</u>: If z is an admissible control (for P_2) with admissible trajectory x and terminal time t_1 then $u = \psi \circ z$ is an admissible control (for P_1). Furthermore, x is the (unique) trajectory and t_1 the (least) terminal time associated with u and $F_1(u) = F_2(z)$.

<u>Proof</u>: From lemma 3.1 the function $u = \psi \circ z: [t_0, t_1] \to U$ is measurable. By hypothesis $\dot{x}(t) = f(x(t), z(t))$ a.e. on $[t_0, t_1]$. Since $f(x(t), z(t)) = g(x(t), \psi(z(t))) = g(x(t), u(t))$ and since $x(t_0) = a$, $x(t_1) \in S$ it follows that $x: [t_0, t_1] \to X$ is the (unique) admissible trajectory for u. Furthermore, t_1 is the least terminal time for u - for if not then $x(t) \in S$ for some $t \in [t_0, t_1)$ which contradicts the fact that t_1 is the least terminal time for z. Therefore $(u; x, t_1) \in \Omega_1$.

Finally,

$$F_2(z) = \int_{t_0}^{t_1} f_0(x(\tau), z(\tau)) \, d\tau$$

$$= \int_{t_0}^{t_1} g_0(x(\tau), \psi(z(\tau))) \, d\tau$$

$$= \int_{t_0}^{t_1} g_0(x(\tau), u(\tau)) \, d\tau$$

$$= F_1(u).$$

<u>Theorem 3.2</u>: If $(u; x, t_1) \in \Omega_1$ and if there exists a bounded measurable function $z: [t_0, t_1] \to Z$ such that $\psi \circ z = u$ then $(z; x, t_1) \in \Omega_2$ and $F_2(z) = F_1(u)$.

Proof: As in theorem 3.1 we obtain $\dot{x}(t) = g(x(t), u(t)) = f(x(t), z(t))$ a.e. and as before we can argue that $(z; x, t_1) \in \Omega_2$. Similarly, we obtain $F_2(z) = F_1(u)$.

Although it is not needed an even stronger result than theorems 3.1 and 3.2 follows from the same type of argument, the definition of the least terminal time and the uniqueness of the trajectories.

Corollary 3.1: If $(u; x, t_1) \in \Omega_1$ and $(z; x^*, t_1^*) \in \Omega_2$ and if $u(t) = \psi(z(t))$ on $[t_0, T]$ for $T = \min \{t_1, t_1^*\}$ then $t_1 = t_1^*$, $x = x^*$, and $F_1(u) = F_2(z)$.

3.3 P_2 is Ψ-Equivalent to P_1

As indicated by theorem 3.2 and pointed out in chapter 1 the key result which enables us to prove P_2 is Ψ-equivalent to P_1 is Filippov's implicit function lemma - a lemma which (for several reasons) is of fundamental importance in the theory of optimal control. With this lemma in mind we make the following definition.

Definition 3.6: The system (ψ, Z, U, R) satisfies the measurable IFP (implicit function property) if for each interval $[t_0, t_1] \subset R$ and (bounded) measurable function $u:[t_0, t_1] \to U$ there exists a bounded, measurable function $z:[t_0, t_1] \to Z$ such that $u = \psi \circ z$.

With regard to the statement made prior to remark 3.4 we have the following.

Corollary 3.2: If (ψ, Z, U, R) satisfies the measurable IFP then $\Omega_2 \neq \emptyset$.

Proof: By assumption $\Omega_1 \neq \emptyset$ and thus there exists a $(u; x, t_1) \in \Omega_1$. By hypothesis there exists a bounded measurable function $z:[t_0, t_1] \to Z$ such that $\psi \circ z = u$ and from theorem 3.2 $(z; x, t_1) \in \Omega_2$.

Prior to remark 3.5 there was some discussion concerning the construction of a function $\Psi: \Omega_2 \to \Omega_1$. We formally do this now, noting that theorem 3.1 justifies the definition.

Definition 3.7: Let $\Psi: \Omega_2 \to \Omega_1$ be defined as $\Psi(z) = u$ if $u = \psi \circ z$.

Corollary 3.3: If (ψ, Z, U, R) satisfies the measurable IFP then Ψ maps Ω_2 onto Ω_1 (i.e., $\Psi(\Omega_2) = \Omega_1$), and in addition $F_2 = F_1 \circ \Psi$.

Proof: That Ψ is onto follows immediately from the arguments of corollary 3.2. Furthermore, if $z \in \Omega_2$ then from theorem 3.1

$$F_2(z) = F_1(\psi \circ z) = F_1(u) \quad (\underline{i.e.,} \quad F_2 = F_1 \circ \Psi).$$

Combining corollaries 3.2 and 3.3 and interpreting corollary 2.2 in terms of the notation of this chapter produces theorem 3.3.

Theorem 3.3: If (ψ, Z, U, R) satisfies the measurable IFP then P_2 is Ψ-equivalent to P_1. In particular, the following statement and its converse are true: for each optimal element $(u^o; x^o, t^o_1) \in \Omega^o$, there exists an optimal element $(z^o; \tilde{x}^o, \tilde{t}^o_1) \in \Omega^o_2$ with $\psi \circ z^o = u^o$. Furthermore, $t^o_1 = \tilde{t}^o_1$, $x^o = \tilde{x}^o$ and $F_1(u^o) = F_2(z^o)$.

3.4 Filippov's Lemma

Definition 3.6 can be stated diagrammatically in the following form. The system (ψ, Z, U, R) satisfies the measurable IFP if there exists a bounded, measurable function z which completes the diagram
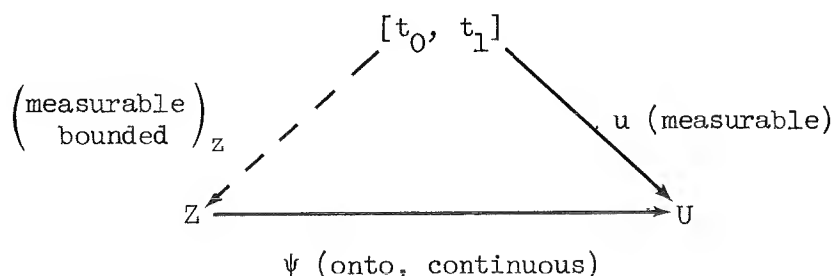

Figure 3.1- Commutative diagram.

If Z is compact and if ψ is one-to-one then ψ is a homeomorphism (i.e., ψ^{-1} exists and is continuous). In this case, given u, we have $z = \psi^{-1} \circ u$ which by lemma 3.1 is measurable and bounded and hence (ψ, Z, U, R) satisfies the measurable IFP. It is this case, with the additional assumption that ψ be C^1, which would normally be considered in a change of variables discussion. However, Filippov's lemma shows that these assumptions are far more restrictive than necessary.

Rather than prove Filippov's lemma completely, we refer the reader to the original paper (5) pages 78-79, or to an identical proof on pages 30-31 of (7). Since we will make a later comment concerning the construction of the function z we discuss this aspect of the lemma. If $K \subset R^p$ is a nonempty compact set, then the continuous function $\alpha_i : K \to R$ where $\alpha_i(z) = z_i$ (i.e., $z = (z_1, \ldots, z_p)$) has a minimum on K i.e., there exists at least one point of K whose i^{th} component is least. Thus by applying α_1 to K we determine an element z^1 whose first component is least. If there is more than one such element, we apply α_2 to the compact set $K \cap \alpha_1^{-1}(\alpha_1(z^1))$ thereby determining from among the set of elements whose first component is least, an element whose second component is least. Again, if there is more than one such element we continue inductively for $i = 1, 2, \ldots, p$ and in this manner uniquely determine a point in K.

Lemma 3.6 (Filippov): Under the conditions of the basic assumption, namely that U is compact, and ψ is continuous with $\psi(Z) = U$ and with the additional assumption that Z is compact, the system (ψ, Z, U, R) satisfies the measurable IFP.

<u>Proof</u>: Let $u:[t_0, t_1] \to U$ be a measurable function and for each $t \in [t_0, t_1]$ let $K(t) = \psi^{-1}(u(t))$. Since $K(t) \subset Z$ and since ψ is continuous the set $K(t)$ is compact and nonempty (ψ is onto). Let $z(t)$ be the unique element of $K(t)$ determined as in the previous discussion and in this manner we define a bounded function $z:[t_0, t_1] \to Z$ such that $\psi \circ z = u$. A proof that z is measurable is based on Luzin's theorem and is contained in the previously mentioned references.

Theorem 3.3 and lemma 3.6 together state that the basic assumption plus the compactness of Z are sufficient for Ψ-equivalence. Frequently (as in chapter 4) it is desirable to remove the compactness assumption on Z and one way of doing this is by strengthening the assumptions on ψ. Doing this one obtains the fundamental result of chapter 3 which is the following.

<u>Theorem 3.4</u>: If the conditions of the basic assumption hold (Z need not be compact) and if in addition there exists a compact set $Z^* \subset Z$ such that ψ maps Z^* onto U then P_2 is Ψ-equivalent to P_1. Furthermore, if $(z; x, t_1) \in \Omega_2$ then there exists a measurable function $z^*:[t_0, t_1] \to Z^*$ with $(z^*; x, t_1) \in \Omega_2$ such that $\psi \circ z = \psi \circ z^*$ and $F_2(z) = F_2(z^*)$. Finally, if $(z; x, t_1) \in \Omega_2^0$ (<u>i.e.</u>, z is an optimal control) then $(z^*; x, t_1) \in \Omega_2^0$.

<u>Proof</u>: By lemma 3.6 the system $(\psi|_{Z^*}, Z^*, U, R)^1$ satisfies the measurable IFP so that (ψ, Z, U, R) does also since $Z^* \subset Z$. From theorem 3.3 it follows that P_2 is Ψ-equivalent to P_1.

Now if $(z; x, t_1) \in \Omega_2$ then $(u; x, t_1) \in \Omega_1$ where $u = \psi \circ z$. Since $(\psi|_{Z^*}, Z^*, U, R)$ satisfies the measurable IFP there exists a (bounded) measurable function $z^*:[t_0, t_1] \to Z$ with $\psi \circ z^* = u$. From theorem 3.2 $(z^*; x, t_1) \in \Omega_2$

---

[1]The function $\psi|_{Z^*}$ denotes the <u>restriction</u> of $\psi: Z \to U$ to $Z^* \subset Z$

and $F_2(z^*) = F_1(u) = F_2(z)$. That the functions $\psi \circ z$ and $\psi \circ z^*$ are equal follows since both are equal to u. Finally, if $(z; x, t_1) \in \Omega^o_2$ then since $F_2(z) = F_2(z^*)$ we have $(z^*; x, t_1) \in \Omega^o_2$ and theorem 3.4 is established.

Paraphrasing the results of the previous theorem, one finds that, under the hypothesis of theorem 3.4, P_2 is Ψ-equivalent to P_1 but without loss of generality one can select the optimal controls for P_2 (if any exist) from among the set of (bounded) measurable functions whose range is contained in Z^*.

Before closing this chapter, some observations are appropriate. First, if the hypotheses of theorem 3.4 are satisfied with $Z = R^p$ then P_1 has been transformed into a problem P_2 whose control region is the whole space R^p. As is illustrated in chapter 4, this provides a bridge between the calculus of variations and the theory of optimal control i.e., a direct method is available for relating the results of the calculus of variation, (e.g., sufficiency theorems, Hamilton-Jacobi theory and, of course, necessary conditions) to problems of optimal control. In addition, since many numerical algorithms are based on the (implicit or explicit) assumption that the control region is open, this provides a method of applying these algorithms in the case where U is compact.

Second, as indicated in (6), there may be some advantage to replacing a problem whose control region has corners by one whose control region is smooth (in some sense). From theorem 3.3 (or 3.4 with $Z = Z^*$) if we may choose Z as, for example, the closed unit ball in R^p this goal is accomplished. We will say more about this in chapter 6.

Third, Filippov's lemma recently has been the subject of research, see for example (12), (16). Particularly the paper of McShane and Warfield (16) generalizes the lemma to include the case where Z is a separable metric space, U is a Hausdorff space, and $[t_0, t_1]$ becomes a subset of a measure space. Thus, in our case, it appears that we could relax the hypotheses of theorems 3.3 and 3.4. However, since the original version (5) will prove to be completely adequate for use herein, and since the compactness of Z (or Z^*) permits an easy visualization of the construction of $z:[t_0, t_1] \rightarrow Z$ we choose to use Filippov's lemma in its original form. In fact, lemma 3.6 is actually less general than the original and is an immediate corollary of it. We remark that the original was intended for use in an existence theorem and that its application herein serves to underscore the fundamental role this lemma plays in the theory of optimal control. (See the discussion on pages 293-297 of (22)).

Finally, loosely speaking, all of what was done in this chapter could be interpreted in terms of piecewise continuous (as opposed to bounded, measurable) controls. However, to do this it would be necessary to obtain the piecewise continuous analogue of Filippov's lemma (i.e., if u is piecewise continuous under what additional assumptions (if any) on U, Z, and ψ can we assume that z is piecewise continuous).

4. LAGRANGE PROBLEMS AS OPTIMAL CONTROL PROBLEMS

On several occasions in chapter 3 it was mentioned that we would be interested in transforming a problem P_1, with a compact control region, into a problem P_2 with R^p as its control region. This is the idea which will be pursued in this chapter and with this in mind, let us strengthen the basic assumption of chapter 3 to read: the set Z is all of R^p and there exists a compact subset $Z^* \subset Z$ such that ψ maps Z^* onto U where U is compact and connected. We remark that since $Z = R^p$ is connected and $\psi(Z) = U$, then by necessity we must assume U is connected. This set of assumptions will be referred to as the strengthened basic assumption.

Remark 4.1: If the strengthened basic assumption is satisfied then theorem 3.4 applies.

In chapter 6 we show that for certain types of compact, connected sets U it is possible to find a positive integer p, continuous function ψ, and compact set Z^* which satisfy the strengthened basic assumption. These results combined with remark 4.1 (i.e., theorem 3.4) verify that for a rather arbitrary control region U we may transform P_1 into a problem P_2 which is ψ-equivalent to P_1 and which has an unconstrained control region. While in many instances this may be adequate, if one wants to cast problem P_2 into the traditional framework of the calculus of variations some additional assumptions are necessary. We list these in the following paragraph.

Throughout the remainder of this chapter assume that the strengthened basic assumption is satisfied and that in addition the function ψ is C^1 on $Z = R^p$. Furthermore, suppose there exists an open set $V \subset R^m$ with $U \subset V$

such that the functions g_0 and $g = (g_1, \ldots g_n)$ are c^1 on $X \times V$. With these additional assumptions the following is true.

Lemma 4.1: The functions f_0 and $f = (f_1, \ldots, f_n)$ (definition 3.5) are c^1 on $X \times Z$ (i.e., on R^{n+p}).

For convenience let us restate problem P_2 in the form it will assume throughout the remainder of this chapter.

Problem P_2: To find a bounded, measurable control function $z^0 = (z^0_1, \ldots, z^0_p)$ with $z^0(t) \in Z = R^p$, a corresponding (unique) absolutely continuous trajectory $x^0 = (x^0_1, \ldots, x^0_n)$ and a (first) terminal time $t^0_1 > t_0$ such that

(1) $x^0(t_0) = a$

(2) $x^0(t^0_1) \in S$

(3) $\dot{x}^0(t) = f(x^0_1(t), \ldots, x^0_n(t), z^0_1(t), \ldots, z^0_p(t))$ a.e. on $[t_0, t^0_1]$

and such that $(z^0; x^0, t^0_1)$ minimizes $F_2(z) = \int_{t_0}^{t_1} f_0(x(\tau), z(\tau)) d\tau$ with respect to all function x, z and terminal times t_1 satisfying these conditions.

Since $X \times Z = R^{n+p}$ and since the functions f_0, f are c^1 on $X \times Z$ it is well known (see (18), chapter V) that P_2 is "equivalent" to a certain Lagrange problem in the calculus of variations. We show herein that this Lagrange problem is in fact Ψ-equivalent to P_2. Since one of the primary objectives of this paper is to illustrate how one can transform an optimal control problem into a calculus of variation problem - thereby correlating known results in the two fields - we will go into some detail in the following discussion.

For arbitrary vectors $y = (y_1, \ldots, y_{n+p})$ and $\dot{y} = (\dot{y}_1, \ldots, \dot{y}_{n+p})$ in R^{n+p} define functions

$$\theta_i(y, \dot{y}) = f_i(y_1, \ldots, y_n, \dot{y}_{n+1}, \ldots, \dot{y}_{n+p})$$

for $i = 0, 1, 2, \ldots, n$ and

$$\varphi_i(y, \dot{y}) = \dot{y}_i - \theta_i(y, \dot{y})$$

for $i = 1, 2, \ldots, n$.

(Later we will introduce a function y and its derivative $\dot{y}$ which are not to be confused with the arbitrary vectors y and $\dot{y}$ used here.) The next remark follows immediately.

Remark 4.2: The functions θ_0, $\theta = (\theta_1, \ldots, \theta_n)$ and $\varphi = (\varphi_1, \ldots, \varphi_n)$ are C^1 on $R^{n+p} \times R^{n+p}$.

In terms of the functions θ_0 and φ let us define a Lagrange problem and then prove that it is Ψ-equivalent to P_2. Since this is the only Lagrange problem considered in this paper, it will be referred to as the Lagrange problem or sometimes as problem P_3. Consistent with the notation of chapter 2 and the terminology of the calculus of variations let Y_3 denote the set of all absolutely continuous functions $y = (y_1, \ldots, y_{n+p})$: $[t_0, t_*] \to R^{n+p}$ for some $t_* > t_0$ such that

$$y(t) = \int_{t_0}^{t} w(\tau) \, d\tau + y^0$$

for an arbitrary vector $y^0 \in R^{n+p}$ and bounded, measurable function
$w:[t_0, t_*] \to R^{n+p}$. By convention we agree to call the function w the
derivative of y and in this way extend the definition of $\dot{y}$ to all of
$[t_0, t_*]$ i.e., if $y \in Y_3$ then $\frac{d}{dt} y$ exists a.e. and $\dot{y} = \frac{d}{dt} y$ a.e. for
a bounded measurable function $w:[t_0, t_*] \to R^{n+p}$ with $w = \dot{y}$.

Definition 4.1: If corresponding to $t_1 > t_0$ there exists a function
$y:[t_0, t_1] \to R^{n+p}$ in Y_3 such that

$$0 = \varphi(y(t), \dot{y}(t))$$

a.e.,

$$y(t_0) = (a_1, \ldots, a_n, 0, \ldots, 0)$$

and

$$y(t_1) \in S \times R^p$$

then y is said to be an admissible arc and t_1 is a terminal time.

As in lemma 3.2 we can assume without loss of generality that t_1 is
the least terminal time and, as in lemma 3.5, that the functional

$$F_3(y) = \int_{t_0}^{t_1} \theta_0(y(\tau), \dot{y}(\tau)) \, d\tau$$

is defined for each admissible arc. Let $\Omega_3 \subset Y_3$ denote the set of admissible
arcs so that F_3 is a functional on Ω_3 and problem P_3 becomes: find a
$y^0 \in \Omega_3$ such that $F_3(y^0) \leq F_3(y)$ for all $y \in \Omega_3$. If such a y^0 exists
we call it a minimal admissible arc and we let Ω^0_3 denote the set of all
such arcs. As noted in chapter 2, P_3 is a global problem and since much of
the calculus of variations is concerned with local (as opposed to global)
minimal arcs the global nature of P_3 should be kept in mind to avoid possible
confusion.

If $y \in \Omega_3$ with terminal time t_1 define functions $x:[t_0, t_1] \to R^n$ and $z:[t_0, t_1] \to R^p$ by $x_i = y_i$ for $i = 1, 2, \ldots, n$ and $z_i = \dot{y}_{n+i}$ for $i = 1, 2, \ldots, p$. Since $\varphi(y(t), \dot{y}(t)) = 0$ a.e. it follows from the definition of φ that $\dot{x}(t) = f(x(t), z(t))$ a.e. and that $(z; x, t_1) \in \Omega_2$. Also from the definition of θ_0 we have $F_3(y) = F_2(z)$. Conversely starting with $(z; x, t_1) \in \Omega_2$ the function $y:[t_0, t_1] \to R^{n+p}$ defined by

$$y_i = x_i$$

for $i = 1, 2, \ldots, n$ and

$$y_{n+i}(t) = \int_{t_0}^{t} z_i(\tau) \, d\tau$$

for $i = 1, 2, \ldots, p$

is in Ω_3. Since $\Omega_2 \neq \emptyset$ we have $\Omega_3 \neq \emptyset$ and in addition the following lemma is true.

Lemma 4.2: The function Ψ defined by $\Psi(y) = z$ where $y_i = x_i$ for $i = 1, 2, \ldots, n$ and $\dot{y}_{n+i} = z_i$ for $i = 1, 2, \ldots, p$ maps Ω_3 onto Ω_2 and $F_3 = F_2 \circ \Psi$.

Therefore, in the terminology of this chapter, corollary 2.2 becomes the following.

Theorem 4.1: The Lagrange problem (P_3) is Ψ-equivalent[1] to P_2. In particular, the following statement and its converse are true: for each minimal admissible arc y^0 and terminal time t_1 there exists an optimal control z^0, optimal trajectory x^0 and terminal time t^*_1. Furthermore,

───────────────

[1] In fact Ψ is one-to-one and hence P_2, P_3 are strictly equivalent in the sense of definition 2.2.

these quantities are related by $y^o_i = x^o_i$ for $i = 1, 2, \ldots, n$

$\dot{y}^o_{n+i} = z^o_i$ for $i = 1, 2, \ldots, p$ and $t_1 = t^*_1$.

Remark 4.2: From remark 2.2, the Lagrange problem is $\Psi_1 \circ \Psi_2$-equivalent

to P_1 where Ψ_1 is defined in definition 3.7 and Ψ_2 is defined in

lemma 4.2.

In chapter 5 we will derive the (Pontryagin) maximum principle for P_2

using the necessary conditions of the calculus of variations and theorem 4.1.

From this we obtain the maximum principle for P_1 using theorem 3.4. We

remark that as a consequence of remark 4.2 we could skip the intermediate

step (P_2) - however, we choose not to do this.

5. AN APPLICATION OF Ψ-EQUIVALENCE

In this chapter we will apply the necessary conditions of the calculus of variations to the Lagrange problem (P_3) formulated in chapter 4 and, using the fact that P_3 is Ψ-equivalent to P_2, derive the (Pontryagin) maximum principle for P_2. Furthermore, by using the fact that P_2 is Ψ-equivalent to P_1, we ultimately obtain the maximum principle for P_1. This result (i.e., a proof of the maximum principle for the original optimal control problem) should not be interpreted as an implication that the material of chapters 2 through 4 is no more than preliminary to the material of chapter 5. On the contrary, chapter 5 is included in this paper because it contains an immediate application of the results of chapters 2, 3, 4, and 6 and because the material of this chapter serves to indicate the basic foundations required for a study of the correspondences between the calculus of variations and optimal control theory.

5.1 Necessary Conditions From the Calculus of Variations

In chapter V of (18), Pontryagin et al. consider a Lagrange problem and, in effect, by transforming the Lagrange problem into an optimal control problem they obtain the Weierstrass condition and multiplier rule from the maximum principle. Thus, their discussion is parallel, but opposite in direction, to the initial developments of this chapter (namely, to the derivation of the maximum principle for P_2, starting from P_3). This paragraph has been included solely for the convenience of the reader who, if he consults (18), should be careful to observe that (in our notation) the Lagrange problem of Pontryagin et al. is slightly more general than the

Lagrange problem considered herein. Specifically, they permit the functions θ_0, θ_1, ..., θ_n to be dependent not only on y_1, ..., y_n but also on y_{n+1}, ..., y_{n+p}. Furthermore, θ_0 may be explicitly dependent on $\dot{y}_1$, ..., $\dot{y}_n$.

For a long time the standard reference on the classical problem of Lagrange has been part II of (1). A more modern introduction to the (classical) theory is contained in (20), which, incidentally, contains a discussion (pages 316-323) similar in content to the discussion of (18), chapter V. Also (9) contains a great deal of information on the Lagrange problem and its relation to the theory of optimal control.[1] However, all of these references, and in addition, many others (with the notable exception of (22)), discuss the Lagrange problem with the assumption that the trajectories are at least piecewise smooth, rather than absolutely continuous. Furthermore, in order to elevate the Weierstrass condition, multiplier rule and transversality conditions to the same level of generality as the maximum principle these necessary conditions must be established without any assumptions regarding normality. For these two reasons, we turn to the work of E. J. McShane and specifically to (15), page 24, for a statement of the previously mentioned necessary conditions. Note, however, that this reference deals with necessary conditions for a parametric problem and consequently, since P_2 is nonparametric, the results must be transformed to results for a nonparametric problem.

---

[1]In contrast to the transformation method outlined herein, (9) uses the method of "slack variables" - usually attributed to Valentine (21) - to bridge the gap between optimal control theory and the calculus of variations.

With these introductory remarks disposed of, let us turn to a discussion

of P_2 and the equivalent Lagrange problem. Since a discussion of the

structure of the target set S would only serve to confuse the fundamental

results, we let S be a single point, say $b \in R^n$, with $a \neq b = (b_1, \ldots, b_n)$.

Therefore, in the traditional spirit of the calculus of variations the terminal

condition t_1 free and $y(t_1) \in \{b\} \times R^p$ would be written as $\sigma(t_1, y(t_1)) = 0$

where the function $\sigma = (\sigma_1, \ldots, \sigma_n)$ is defined by $\sigma_i(t, y) = y_i - b_i$ for

$i = 1, 2, \ldots, n$. (i.e., $\sigma = 0$ defines the closed terminal manifold

$[t_0, \infty) \times \{b\} \times R^p$ in (t, y)- space.) For convenience, we state the Lagrange

problem considered in this chapter.

Lagrange Problem: For fixed initial conditions $(a, t_0) \in R^{n+1}$ and terminal

conditions $\sigma = 0$ find a (first) time $t_1 > t_0$ and function $y = (y_1, \ldots, y_{n+p})$

in Y_3 such that[1]

(1) $y(t_0) = (a_1, \ldots, a_n, 0, \ldots, 0)$

(2) $\sigma(t_1, y(t_1)) = 0$

(3) $\varphi(y(t), \dot{y}(t)) = 0$ a.e.

and such that y minimizes

$$F_3(y) = \int_{t_0}^{t_1} \theta_0(y(\tau), \dot{y}(\tau)) \, d\tau$$

with respect to all other such functions (i.e., with respect to all admissible

arcs).

_____

[1]Recall our earlier (chapter 4) convention regarding Y_3.

In chapter 4 we remarked that the functions θ_0, θ and φ are C^1 on $R^{n+p} \times R^{n+p}$. The following lemma is an immediate consequence of the definitions of these functions and is merely an expression of the fact that θ is independent of $y_{n+1}, \ldots, y_{n+p}$ and $\dot{y}_1, \ldots, \dot{y}_n$.

Lemma 5.1: For all $(y, \dot{y}) \in R^{n+p} \times R^{n+p}$ the following is true:

(1) $\dfrac{\partial \varphi_i}{\partial y_j} = - \dfrac{\partial \theta_i}{\partial y_j} \; ; \; \dfrac{\partial \varphi_i}{\partial \dot{y}_j} = \delta_{ij} ; \; \dfrac{\partial \theta_i}{\partial \dot{y}_j} = 0$ $\begin{array}{l} i = 1, 2, \ldots, n \\ j = 1, 2, \ldots, n \end{array}$

(2) $\dfrac{\partial \varphi_i}{\partial y_j} = - \dfrac{\partial \theta_i}{\partial y_j} = 0; \; \dfrac{\partial \varphi_i}{\partial \dot{y}_j} = - \dfrac{\partial \theta_i}{\partial \dot{y}_j}$ $\begin{array}{l} i = 1, 2, \ldots, n \\ j = n+1, \ldots, n+p \end{array}$

(3) $\text{rank} \left(\dfrac{\partial \varphi_i}{\partial \dot{y}_j} \right) = n$ where $\left(\dfrac{\partial \varphi_i}{\partial \dot{y}_j} \right)$ denotes the $n \times (n+p)$ matrix with elements $\dfrac{\partial \varphi_i}{\partial \dot{y}_j}$.

In anticipation of what will be needed later let us define, for arbitrary vectors $\lambda = (\lambda_0, \ldots, \lambda_n) \in R^{n+1}$, $\xi = (\xi_1, \ldots, \xi_{n+p}) \in R^{n+p}$ and $(y, \dot{y}) \in R^{n+p} \times R^{n+p}$ the (Lagrange) function

$$h(y, \dot{y}, \lambda) = - \lambda_0 \theta_0(y, \dot{y}) + \sum_{i=1}^{n} \lambda_i \varphi_i(y, \dot{y}),$$

the (Weierstrass) excess function

$$E(y, \dot{y}, \xi, \lambda) = h(y, \xi, \lambda) - h(y, \dot{y}, \lambda) - \sum_{i=1}^{n+p} (\xi_i - \dot{y}_i) \frac{\partial h}{\partial \dot{y}_i} (y, \dot{y}, \lambda),$$

and the (Pontryagin) function

$$G(y, \dot{y}, \lambda) = \sum_{i=0}^{n} \lambda_i \theta_i(y, \dot{y}).$$

We note that the functions h and G are C^1 on $R^{n+p} \times R^{n+p} \times R^{n+1}$.

From the calculus of variations we obtain the following necessary conditions for y to be a solution to the Lagrange problem. We designate it as theorem 5.1 and accept it as being proved. It is from this theorem that the maximum principle for both P_2 and P_1 will be derived.

Theorem 5.1: If y is a minimal admissible arc with terminal time t_1 for the Lagrange problem (i.e., $(y; t_1) \in \Omega^0_3$), then there exists a nonzero, bounded, measurable function $\lambda^* = (\lambda^*_0, \ldots, \lambda^*_n)$ on $[t_0, t_1]$ with $\lambda^*_0 \leq 0$ a constant, such that for almost all $t \in [t_0, t_1]$:

(1) $\quad \dfrac{\partial h}{\partial \dot{y}_i} (y(t), \dot{y}(t), \lambda^*(t)) = \displaystyle\int_{t_0}^{t} \dfrac{\partial h}{\partial y_i} (y(s), \dot{y}(s), \lambda^*(s)) \, ds + C_i$

for constants C_i where $i = 1, 2, \ldots, n+p$,

(2) $\quad h(y(t), \dot{y}(t), \lambda^*(t)) - \displaystyle\sum_{i=1}^{n+p} \dot{y}_i(t) \dfrac{\partial h}{\partial \dot{y}_i} (y(t), \dot{y}(t), \lambda^*(t))$

$$= \int_{t_0}^{t} \frac{\partial h}{\partial t} (y(s), \dot{y}(s), \lambda^*(s)) \, ds + C_0$$

for a constant C_0, and

(3) $\quad E(y(t), \dot{y}(t), \xi, \lambda^*(t)) \geq 0 \qquad$ for all vectors ξ satisfying $\varphi(y(t), \xi) = 0$.

In addition at the terminal time $t = t_1$ it is necessary that there exist constants $\nu_1, \ldots, \nu_n$ with $(\lambda_0, \nu_1, \ldots, \nu_n) \neq 0$ such that

$$\sum_{j=1}^{n} v_j \frac{\partial \sigma_j}{\partial t} (t_1, y(t_1)) + C_0 + \int_{t_0}^{t_1} \frac{\partial h}{\partial t} (y(s), \dot{y}(s), \lambda^*(s)) = 0$$

and

$$\sum_{j=1}^{n} v_j \frac{\partial \sigma_j}{\partial y_i} (t_1, y(t_1)) + C_i + \int_{t_0}^{t_1} \frac{\partial h}{\partial y_i} (y(s), \dot{y}(s), \lambda^*(s)) \, ds = 0$$

where $i = 1, 2, \ldots, n+p$.

Theorem 5.1 is stated in full generality and, as yet, makes no use of the particular form h, E and σ assume herein. At this time we choose to make use of this special form and, in the spirit of optimal control theory, interpret theorem 5.1 in terms of the function G defined previously. Before doing this, let us, for convenience, state the following lemma which is merely an algebraic consequence of lemma 5.1 and the relation

$$h(y, \dot{y}, \lambda) = \sum_{i=1}^{n} \lambda_i \dot{y}_i - G(y, \dot{y}, \lambda).$$

<u>Lemma 5.2</u>: For all $y, \dot{y} \in R^{n+p}$ and $\lambda \in R^{n+1}$

(1) $\dfrac{\partial h}{\partial y_i} = - \dfrac{\partial G}{\partial y_i} \; ; \; \dfrac{\partial h}{\partial \dot{y}_i} = \lambda_i \; ; \; \dfrac{\partial G}{\partial \dot{y}_i} = 0$ $i = 1, 2, \ldots, n$

(2) $\dfrac{\partial h}{\partial y_i} = - \dfrac{\partial G}{\partial y_i} = 0 \; ; \; \dfrac{\partial h}{\partial \dot{y}_i} = - \dfrac{\partial G}{\partial \dot{y}_i}$ $i = n+1, \ldots, n+p$

(3) $h - \displaystyle\sum_{i=1}^{n+p} \dot{y}_i \frac{\partial h}{\partial \dot{y}_i} = - G + \sum_{i=n+1}^{n+p} \dot{y}_i \frac{\partial G}{\partial \dot{y}_i}$

and for arbitrary $\xi \in R^{n+p}$

$$(4) \quad E(y, \dot{y}, \xi, \lambda) = G(y, \dot{y}, \lambda) - G(y, \xi, \lambda)$$

$$+ \sum_{i=n+1}^{n+p} (\xi_i - \dot{y}_i) \frac{\partial G}{\partial \dot{y}_i} (y, \dot{y}, \lambda).$$

If y is a minimal admissable arc for the Lagrange problem, then there exists a constant $\lambda^*_0 \leq 0$ and a function $\lambda^* = (\lambda^*_0, \lambda^*_1, \ldots, \lambda^*_n)$ satisfying the conditions of theorem 5.1. Since λ^* is bounded, measurable and $\frac{\partial G}{\partial y_i}$ is continuous we have, from lemma 3.1, that $\frac{\partial G}{\partial y_i} (y(t), \dot{y}(t), \lambda^*(t))$ is integrable and hence the functions λ_i, $i = 1, 2, \ldots, n$ given by

$$\lambda_i(t) = - \int_{t_0}^{t} \frac{\partial G}{\partial y_i} (y(s), \dot{y}(s), \lambda^*(s)) \, ds + C_i$$

are absolutely continuous and satisfy $\dot{\lambda}_i(t) = - \frac{\partial G}{\partial y_i} (y(t), \dot{y}(t), \lambda^*(t))$ a.e. on $[t_0, t_1]$. From (1) of theorem 5.1 and (1) of lemma 5.2 it follows that $\lambda_i = \lambda^*_i$ a.e. and thus, letting $\lambda_0 = \lambda^*_0$ we have that $\lambda_0 \leq 0$, the function $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_n)$ is nonzero and absolutely continuous, and $\dot{\lambda}_i(t) = - \frac{\partial G}{\partial y_i} (y(t), \dot{y}(t), \lambda(t))$ a.e.

Turning to the terminal conditions of theorem 5.1 we find that, since

$$\frac{\partial \sigma_i}{\partial t} = \frac{\partial h}{\partial t} = \frac{\partial \sigma_i}{\partial y_j} = \frac{\partial h}{\partial y_j} = 0$$

for   i = 1, 2, ..., n   and   j = n+1, ..., n+p

the constants C_0 and C_{n+1}, ..., C_{n+p} are zero. Therefore from (2) of

lemma 5.2 and (1) of theorem 5.1, plus that fact that $\lambda = \lambda^*$ a.e., we

obtain

$$0 = \frac{\partial G}{\partial \dot{y}_i}\,(y(t),\,\dot{y}(t),\,\lambda(t))$$

a.e. for   i = n+1, ..., n+p.

From (2) of theorem 5.1 and (3) of lemma 5.2 we have that

$$0 = G(y(t),\,\dot{y}(t),\,\lambda^*(t))$$

a.e. on $[t_0,\,t_1]$ and from (4) of lemma 5.2 we have $G(y(t),\,\dot{y}(t),\,\lambda^*(t)) \geq$
$G(y(t),\,\xi,\,\lambda^*(t))$ a.e. for all vectors ξ such that $\varphi(y(t),\,\xi) = 0$. Since
$\lambda = \lambda^*$ a.e. these two relations are also true for almost all $t \in [t_0,\,t_1]$
with λ^* replaced by λ. In other words we have established the following.

Corollary 5.1: If $(y;\,t_1) \in \Omega^o_3$ then there exists a nonzero, absolutely
continuous function $\lambda = (\lambda_0,\,\lambda_1,\,...,\,\lambda_n)$, with $\lambda_0 \leq 0$ a constant, such
that a.e. on $[t_0,\,t_1]$:

(1)' $\dot{\lambda}_i(t) = -\frac{\partial G}{\partial y_i}\,(y(t),\,\dot{y}(t),\,\lambda(t))$ i = 1, 2, ..., n

$\frac{\partial G}{\partial \dot{y}_i}\,(y(t),\,\dot{y}(t),\,\lambda(t)) = 0$ i = n+1, ..., n+p

(2)' $G(y(t), \dot{y}(t), \lambda(t)) = 0$

(3)' $G(y(t), \dot{y}(t), \lambda(t)) \geq G(y(t), \xi, \lambda(t))$

for all vectors ξ satisfying $\varphi(y(t), \xi) = 0$.

5.2 Necessary Conditions for P_2

Turning to problem P_2 (see chapter 4) let us define, for arbitrary vectors $(x, z, \lambda) \in R^n \times R^p \times R^{n+1}$, the (Pontryagin) hamiltonian

$$\overline{H}(x, z, \lambda) = \sum_{i=0}^{n} \lambda_i f_i(x, z).$$

From the properties of f_i we have immediately that $\overline{H}$ is C^1 on $R^n \times R^p \times R^{n+1}$ and furthermore that the following lemma is true.

<u>Lemma 5.3</u>: If the vectors $(y, \dot{y}) \in R^{n+p} \times R^{n+p}$ and $(x, z) \in R^n \times R^p$ satisfy $y_i = x_i$ for $i = 1, 2, \ldots, n$ and $\dot{y}_{n+i} = z_i$ for $i = 1, 2, \ldots, p$ then (independent of $y_{n+1}, \ldots, y_{n+p}$ and $\dot{y}_1, \ldots, \dot{y}_n$) for $\lambda \in R^{n+1}$

(1) $\overline{H}(x, z, \lambda) = G(y, \dot{y}, \lambda)$

(2) $\dfrac{\partial \overline{H}}{\partial x_i}(x, z, \lambda) = \dfrac{\partial G}{\partial y_i}(y, \dot{y}, \lambda)$ \qquad\qquad $i = 1, 2, \ldots, n$

(3) $\dfrac{\partial \overline{H}}{\partial z_i}(x, z, \lambda) = \dfrac{\partial G}{\partial \dot{y}_{n+i}}(y, \dot{y}, \lambda)$ \qquad\qquad $i = 1, 2, \ldots, p$.

From lemma 5.3, the fact that P_3 is Ψ-equivalent to P_2, and corollary 5.1, we obtain the maximum principle for P_2.

<u>Corollary 5.2</u>   (Maximum principle for P_2):   If $z: [t_0, \ t_1] \to Z = R^p$

is an optimal control with corresponding trajectory   x   and terminal time

$t_1 > t_0$ then there exists a nonzero, absolutely continuous function

$\lambda = (\lambda_0, \ \lambda_1, \ \ldots, \ \lambda_n)$ with $\lambda_0 \leq 0$ a constant, such that a.e. on $[t_0, \ t_1]$:

$$(1)" \quad \dot{\lambda}_i(t) = - \frac{\partial \overline{H}}{\partial x_i} \ (x(t), \ z(t), \ \lambda(t)) \qquad i = 1, \ 2, \ \ldots, \ n$$

$$\frac{\partial \overline{H}}{\partial z_i} \ (x(t), \ z(t), \ \lambda(t)) = 0 \qquad i = 1, \ 2, \ \ldots, \ p$$

$$(2)" \quad \overline{H}(x(t), \ z(t), \ \lambda(t)) = 0$$

$$(3)" \quad \overline{H}(x(t), \ z(t), \ \lambda(t)) \geq \overline{H}(x(t), \ \xi, \ \lambda(t))$$

for all vectors $\xi = (\xi_1, \ \ldots, \ \xi_p) \in Z$.

<u>Proof</u>:   If $(z; \ x, \ t_1) \in \Omega^o_2$ then from theorem 4.1 there exists $(y, t_1) \in \Omega^o_3$

with $x_i(t) = y_i(t)$ for $i = 1, \ 2, \ \ldots, \ n$ and $z_i(t) = \dot{y}_{n+i}(t)$ for

$i = 1, \ 2, \ \ldots, \ p$.   Since $(y, \ t_1) \in \Omega^o_3$ there exists a constant $\lambda_0 \leq 0$ and

functions λ_i satisfying corollary 5.1.   Thus from lemma 5.3 relations (1)"

and (2)" follow and furthermore if $\overline{\xi} = (\overline{\xi}_1, \ \ldots, \ \overline{\xi}_{n+p})$ satisfies $\varphi(y(t), \overline{\xi}) = 0$

then the vector $\xi = (\overline{\xi}_{n+1}, \ \ldots, \ \overline{\xi}_{n+p}) \in Z$ satisfies $\overline{H}(x(t), \ \xi, \ \lambda(t))$

$= G(y(t), \ \overline{\xi}, \ \lambda(t))$ which established (3)".

We remark that the relation $\frac{\partial \overline{H}}{\partial z_i} = 0$ follows from (3)" and the fact

that $\overline{H}$ is C^1 on $Z = R^p$.   Therefore, in this sense, the multiplier rule -

when applied to this specific problem (P_3) - contains redundant information.

We turn to problem P_1 and define, for arbitrary vectors $(x, \ u, \ \lambda) \in R^n \times V \times R^{n+1}$

the hamiltonian $H(x, u, \lambda) = \sum_{i=0}^{n} \lambda_i g_i(x, u)$. Again H is C^1 on

$R^n \times V \times R^{n+1}$ and we have the following lemma (recall $X = R^n$ and $U \subset V \subset R^m$).

Lemma 5.4: If the vectors $(x, z) \in X \times Z$ and $(x, u) \in X \times U$ satisfy $\psi(z) = u$ then for $\lambda \in R^{n+1}$

(1) $H(x, u, \lambda) = \overline{H}(x, z, \lambda)$

(2) $\dfrac{\partial H}{\partial x_i}(x, u, \lambda) = \dfrac{\partial \overline{H}}{\partial x_i}(x, z, \lambda)$

(3) $\displaystyle\sum_{j=1}^{m} \dfrac{\partial H}{\partial u_j}(x, u, \lambda)\dfrac{\partial \psi_j}{\partial z_i}(z) = \dfrac{\partial \overline{H}}{\partial z_i}(x, z, \lambda)$, $\quad i = 1, 2, \ldots, p$.

If $(u; x, t_1) \in \Omega_1^O$ then from theorem 3.4 there exists $(z; x, t_1) \in \Omega_2^O$ with $\psi \circ z = u$ i.e., $\psi(z(t)) = u(t)$ for all $t \in [t_0, t_1]$. Applying corollary 5.2 and lemma 5.4 we obtain, as in the proof of corollary 5.2, the maximum principle for P_1.

5.3 The Pontryagin Maximum Principle

Corollary 5.3 (Maximum principle for P_1): If $(u; x, t_1) \in \Omega_1^O$ then there exists a nonzero, absolutely continuous function $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_n)$, with $\lambda_0 \leq 0$ a constant, such that a.e. on $[t_0, t_1]$:

(1)''' $\dot{\lambda}_i(t) = -\dfrac{\partial H}{\partial x_i}(x(t), u(t), \lambda(t))$ $\qquad i = 1, 2, \ldots, n$

(2)''' $H(x(t), u(t), \lambda(t)) = 0$

(3)''' $H(x(t), u(t), \lambda(t)) \geq H(x(t), \xi, \lambda(t))$

for all vectors $\xi = (\xi_1, \ldots, \xi_m) \in U$.

A few observations should be made concerning corollary 5.3. First, we choose not to include the relation

$$\sum_{j=1}^{m} \frac{\partial H}{\partial u_j} (x(t),\ u(t),\ \lambda(t)) \frac{\partial \psi_j(z(t))}{\partial z_i} = 0 \quad \text{a.e.}$$

for $i = 1, 2, \ldots, p$ since, as indicated previously, it can be obtained from condition $(3)''''$ and the definition of H. Second, we see that corollary 5.3 can be considered proved only for those particular control regions U for which the existence of a function ψ satisfying the strengthened basic assumption of chapter 4 is guaranteed. This existence question is considered in the next chapter. Third, the proof of corollary 5.3 (or 5.2) rests on the assumption that g (or f) is C^1 on $V \supset U$ (or Z). This is a slightly stronger assumption than the one made by Pontryagin et al. who only require continuity on U (or Z). Finally, as discussed on pages 101-104 of (18), one can show that the function

$$\alpha(t) = \sup_{\xi \in U} H(x(t),\ \xi,\ \lambda(t))$$

is equal to zero for all $t \in [t_0,\ t_1]$ and not just almost everywhere.

Since a proof of the multiplier role, Weierstrass condition and transversality conditions, without any assumptions regarding normality and with the trajectories absolutely continuous, appears to be approximately as difficult as a proof of the maximum principle and since the proof of the maximum principle outlined in this chapter depends ultimately on Filippov's lemma and the existence of a function ψ satisfying the strengthened basic

assumptions, we do not advocate corollary 5.3 as an "easy" proof of the

maximum principle for an <u>arbitrary</u> U. Rather, we present corollary 5.3

as what it really is - one of (hopefully) many straightforward applications

of the idea of Ψ-equivalence (or what is the same thing, the idea of a

change of variables in the control space).

6. THE EXISTENCE OF ψ

In this the sixth chapter, we turn to the question of the existence of a function ψ satisfying various forms of what has been called (in chapters 3, 4, and 5) the basic assumption. To facilitate later reference, let us systematically list these various forms, in order of decreasing generality, as follows.

Basic Assumption: Corresponding to a nonempty, <u>compact</u> set $U \subset R^m$

(6-1) - there exists a positive integer p, a nonempty set $Z \subset R^p$ and a <u>continuous</u> function $\psi : Z \to U$ with $\psi(Z) = U$;

(6-2) - in addition to (6-1) there exists a <u>compact</u> set $Z^x \subset Z$ such that $\psi(Z^x) = U$;

(6-3) - in addition to (6-2), $Z = R^p$ and (hence) U is <u>connected</u>;

(6-4) - in addition to (6-3), ψ is C^1 on Z.

Recall that, from theorem 3.4, if (6-2) is satisfied then P_2 is ψ-equivalent to P_1. Furthermore, if (6-3) is satisfied then P_2 has an unconstrained control region. Finally, if (6-4) is satisfied then P_2 (with some additional assumptions concerning the function g) is - in the sense of chapter 5 - a Lagrange problem.

Rather than attack the question of the existence of ψ in full generality, let us agree hereafter to consider control regions U which are compact and convex. It will soon be clear that what we lose in generality we will more than gain in practicality. Specifically, when U is compact and convex we will obtain not only existence theorems but also explicit representations for ψ.

Moreover, in most cases the form that ψ assumes will be of sufficient simplicity to guarantee that the transformation $f_i(x, z) = g_i(x, \psi(z))$ is not only possible but also practical. At the end of this chapter we will (briefly) consider a particular type of compact, connected (but not, in general, convex) control region; namely we will let U be the continuous image of a compact, convex set.

6.1 U is a Convex Body

Definition 6.1: For a positive integer p and real numbers a, b with $a < b$, let $[a, b]^p = [a, b] \times \cdots \times [a, b]$ denote a p-dimensional cube while $B^p(\rho, z) = \{x \in R^p : \| x-z \| \leq \rho\}$ denotes[1] the p-dimensional ball of radius $\rho > 0$ centered at $z \in R^p$.

If $U \subset R^m$ is compact and convex and if in addition U contains an interior point then dimension $U = m$ (11) and U is said to be a convex body. If U is a convex body, we will show that there exists a function ψ satisfying (6-3). Moreover, when U is a polyhedron we will construct a rather simple function ψ satisfying (6-4). Finally, when U is a parallelipiped we will construct a very simple function ψ satisfying (6-4) with $p = m$. As a first step in this direction consider the following lemma which may be found (for example) in (3).

Lemma 6.1: If $U \subset R^m$ is a convex body then U is homeomorphic to $B^m(1, 0)$. Consequently, any two convex bodies in R^m are homeomorphic.

From lemma 6.1 it follows that if U is a convex body in R^m and Z^* is any other convex body in R^m then, with $p = m$ and $Z = Z^*$, there exists

---

[1]Recall that for $x \in R^p$, $\|x\| = \left[\displaystyle\sum_{i=1}^{p} (x_i)^2\right]^{\frac{1}{2}}$.

a function ψ satisfying (6-2) i.e., $\psi : Z \to U$ is the homeomorphism whose existence is guaranteed by lemma 6.1. Therefore, we remark in passing that, in the sense of theorem 3.4, when $U \subset R^m$ is a convex body, there is no loss of generality in assuming that U is $B^m(1, 0)$, or $[-1, 1]^m$, or any other convex body in R^m for that matter. Unfortunately, unless the homeomorphism $\psi : Z \to U$ is C^1 (which in general it is not), we can not use lemma 6.1 to construct a function ψ satisfying (6-4). However, the following is true.

Theorem 6.1: If $U \subset R^m$ is a convex body then there exists a function $\psi : R^p \to U$, with $m = p$ and $Z^* = \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2} \right]^m$ which satisfies (6-3) i.e, ψ is continuous and $\psi(R^m) = \psi(Z^*) = U$.

Proof: Since both U and $[-1, 1]^m$ are convex bodies in R^m, from lemma 6.1 there exists a homeomorphism $\varphi : [-1, 1]^m \to U$. Consider the function $\gamma = (\gamma_1, \ldots, \gamma_m) : R^m \to R^m$ defined by

$$\gamma_i(z) = \sin z_i \qquad i = 1, 2, \ldots, m$$

where $z = (z_1, \ldots, z_m)$. Clearly $\gamma(R^m) = \gamma([-\dfrac{\pi}{2}, \dfrac{\pi}{2}]^m) = [-1, 1]^m$ and since γ is continuous the theorem follows with $\psi = \varphi \circ \gamma$.

Certainly, in the proof of theorem 6.1, if the function φ were C^1 then ψ would be also and hence ψ would satisfy (6-4). Also notice the fact that ψ is one-to-one is not used in the proof. These two observations serve to indicate the course we will pursue; namely, we will relax the one-to-one requirement on ψ in order to obtain the condition $\psi \in C^1$. At the same time by assuming that U is a polyhedron we will be able to obtain an explicit representation for ψ.

6.2 U is a Polyhedron

<u>Definition 6.2</u>: For $m \geq 1$ and $r \geq 1$ let w^i for $i = 0, 1, \ldots, r$ be $r+1$ vectors in R^m. The set given by

$$\left\{ u \in R^m : u = \sum_{i=0}^{r} \mu_i w^i, \; \mu_i \geq 0, \sum_{i=0}^{r} \mu_i = 1 \right\}$$

is said to be a <u>polyhedron</u> (or a <u>convex polytope</u>). As defined herein, a polyhedron is the <u>convex hull</u> of its <u>vertices</u> $\{w^0, \ldots, w^r\}$ (3). If $r = m$ and if the set $\{w^1 - w^0, \ldots, w^r - w^0\}$ is <u>linearly independent</u> then the polyhedron is said to be a <u>simplex</u>. (See (10), chapter 5.)

Clearly a polyhedron is compact and convex although, in general, it may or may not be a convex body. Geometrically a simplex is a compact line segment for $m = 1$, a compact triangular region for $m = 2$, a compact solid tetrahedron for $m = 3$, and so forth, and is, for all m, a convex body.

<u>Lemma 6.2</u>: If $U \subset R^m$ is a polyhedron defined by $r+1$ vertices $\{w^0, \ldots, w^r\}$ then there exists a C^1 function $\varphi : R^r \to R^m$ with $\varphi(B^r(1,0)) = U$.

<u>Proof</u>: By hypothesis we have

$$U = \left\{ u \in R^m : u = \sum_{i=0}^{r} \mu_i w^i, \; \mu_i \geq 0, \sum_{i=0}^{r} \mu_i = 1 \right\}.$$

Notice that if $u \in U$ then we have

$$u = \sum_{i=0}^{r} u_i w^i = \mu_0 w^0 + \sum_{i=1}^{r} \mu_i w^i = \left(1 - \sum_{i=0}^{r} \mu_i\right) w^0 + \sum_{i=1}^{r} \mu_i w^i$$

so that

$$u = w^0 + \sum_{i=1}^{r} \mu_i(w^i - w^0).$$

Therefore, U may be written equivalently as

$$\left\{ u \in R^m : u = w^0 + \sum_{i=1}^{r} \mu_i(w^i - w^0), \ \mu_i \geq 0, \sum_{i=1}^{r} \mu_i \leq 1 \right\}.$$

Consider $\varphi : R^r \to R^m$ defined by

$$\varphi(z) = w^0 + \sum_{i=1}^{r} (z_i)^2 (w^i - w^0)$$

where $z = (z_1, \ldots, z_r)$. Certainly φ is C^1 and $\varphi(B^r(1, 0)) \subset U$.

Moreover, if $u \in U$ then u has a representation as $w^0 + \sum_{i=1}^{r} \mu_i(w^i - w^0)$

with $\mu_i \geq 0$ and $\sum_{i=1}^{r} \mu_i \leq 1$. Defining $z = (z_1, \ldots, z_r)$ as $z_i = \sqrt{\mu_i}$

for $i = 1, 2, \ldots, r$ then $z \in B^r(1, 0)$ and $\varphi(z) = u$ $\underline{i.e.}$, $\varphi(B^r(1, 0)) \supset U$.

Notice that φ is definitely not one-to-one. In fact, even when U is a simplex, we can only conclude that $\varphi(z^1) = \varphi(z^2)$ implies $(z^1_i)^2 = (z^2_i)^2$. This follows from the linear independence of the set $\{ w^1 - w^0, \ldots, w^r - w^0 \}$. However, of course, when U is a simplex then φ $\underline{is}$ one-to-one on (for example) the set $B^m(1, 0) \cap \{ z \in R^m : z_i \geq 0, \ i = 1, 2, \ldots, m \}$.

$\underline{\text{Theorem } 6.2}$: If $U \subset R^m$ is a polyhedron defined by the $r+1$ vertices $\{ w^0, \ldots, w^r \}$ then there exists a C^1 function $\psi : R^r \to U$ ($\underline{i.e.}$, $p = r$)

with $Z^* = B^r(1, 0)$ such that $\psi(R^r) = \psi(Z^*) = U$ <u>i.e.</u>, ψ satisfied (6-4). Such a function is given by

$$\psi(z) = w^0 + \frac{\sin^2 \frac{\pi}{2} \|z\|}{\|z\|^2} \sum_{i=1}^{r} (z_i)^2 (w^i - w^0) \qquad z \neq 0$$

$$\psi(0) = w^0 \qquad\qquad\qquad z = 0$$

where $z = (z_1, \ldots, z_r)$.

<u>Proof</u>: Let φ denote the function given by lemma 6.2 and consider $\gamma : R^r \to R^r$ defined by

$$\gamma(z) = \frac{\sin \frac{\pi}{2} \|z\|}{\|z\|} z \qquad z \neq 0$$

$$\gamma(0) = 0 \qquad\qquad z = 0$$

for $z = (z_1, \ldots, z_r)$. Since $\| \gamma(z) \| = \sin \frac{\pi}{2} \|z\| \leq 1$ it follows that $\gamma(R^r) \subset B^r(1, 0)$. Now, if $u \in B^r(1, 0)$ with $\|u\| \neq 0$ define $z = (z_1, \ldots, z_r)$ as

$$z = \frac{2}{\pi} \frac{\sin^{-1} \|u\|}{\|u\|} u$$

where $\sin^{-1}$ is the inverse of $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$. Since $\|z\| = \frac{2}{\pi} \sin^{-1} \|u\|$ we have $\|z\| \leq 1$ and

$$\gamma(z) = \frac{\sin \frac{\pi}{2} \left(\frac{2}{\pi} \sin^{-1} \|u\| \right)}{\frac{2}{\pi} \sin^{-1} \|u\|} \left(\frac{2}{\pi} \frac{\sin^{-1} \|u\|}{\|u\|} u \right) = u .$$

If $\|u\| = 0$ then $u = 0$ and $\gamma(0) = 0$ by definition. Therefore, $\gamma(B^r(1, 0)) \supset B^r(1, 0)$ and since $\gamma(R^r) \supset \gamma(B^r(1, 0))$ this proves that $\gamma(R^r) = \gamma(B^r(1, 0)) = B^r(1, 0)$.

To prove that γ is C^1 on R^r it is sufficient to prove that $\dfrac{\sin \frac{\pi}{2} \|z\|}{\|z\|}$ is C^1. However, the function $\alpha(s) = \dfrac{\sin \frac{\pi}{2} \sqrt{s}}{\sqrt{s}}$ $(s \neq 0)$, $\alpha(0) = \dfrac{\pi}{2}$, $\alpha'(0) = \left(\dfrac{\pi}{2}\right)^3$ is C^1 on R and β is C^1 on R^r where

$$\beta(z) = \|z\|^2 = z_1^2 + \ldots + z_r^2$$

which implies $\alpha \circ \beta(z) = \dfrac{\sin \frac{\pi}{2} \|z\|}{\|z\|}$ is C^1 on R^r. Therefore, the function $\psi = \varphi \circ \gamma$ is C^1 on R^r and satisfies $\psi(R^r) = \psi(B^r(1, 0)) = U$ which proves the theorem.

Remark 6.1: If $U \subset R^m$ is a simplex then theorem 6.2 is true with $p = r = m$ i.e., $\psi : R^m \to U$ and $\psi(R^m) = U$.

Remark 6.2: The proof of theorem 6.2 is certainly true for any C^1 function $\gamma : R^r \to R^r$ which satisfies $\gamma(R^r) = \gamma(Z^*) = B^r(1, 0)$ where $Z^* \subset R^r$ is compact. Whereas theorem 6.2 uses the function $\gamma(z) = \dfrac{\sin \frac{\pi}{2} \|z\|}{\|z\|} z$ we could have chosen, for example, the following function which is suggested by a consideration of r-dimensional spherical coordinates ((13), page 427).

$$\gamma_1^*(z) = \sin^2 z_1 \cos z_2$$
$$\gamma_2^*(z) = \sin^2 z_1 \sin z_2 \cos z_3$$
$$\vdots$$
$$\gamma_{r-1}^*(z) = \sin^2 z_1 \sin z_2 \ldots \sin z_{n-1} \cos z_n$$
$$\gamma_r^*(z) = \sin^2 z_1 \sin z_2 \ldots \sin z_{n-1} \sin z_n$$

with $Z^* = [-\frac{\pi}{2}, \frac{\pi}{2}]^{r-1} \times [-\pi, \pi]$. Certainly $\gamma^* = (\gamma_1^*, \ldots, \gamma_r^*)$ is C^1 and it can be established with an inductive argument that $\gamma^*(R^r) = \gamma^*(z^*) = B^r(1, 0)$. Hence $\psi^* = \varphi \circ \gamma^*$ also satisfies (6-4).

6.3 U is a Parallelipiped

<u>Definition 6.3</u>: For $m \geq 1$ let w^i for $i = 1, 2, \ldots, m$ be m <u>linearly independent</u> vectors in R^m and let $w^0 \in R^m$ be arbitrary. The set

$\left\{ u \in R^m : u = w^0 + \sum_{i=1}^{m} \mu_i w^i, \; 0 \leq \mu_i \leq 1 \right\}$ is said to be a (m-dimensional) <u>parallelipiped</u>.

It follows from the definition that a parallelipiped is compact and convex. Furthermore, a parallelipiped is the convex hull of its 2^m vertices and hence is a polyhedron. Certainly then, theorem 6.2 applies to the case where $U \subset R^m$ is a parallelipiped and thus there exists a function $\psi : R^p \xrightarrow{\text{onto}} U$ satisfying (6-4) with $p = 2^m - 1$. However, we will show (theorem 6.3) that when U is a parallelipiped there exists a function $\psi : R^p \xrightarrow{\text{onto}} U$ satisfying (6-4) with $p = m$ and that ψ assumes a much "simpler" form than the form of the function given by theorem 6.2.

<u>Lemma 6.3</u>: If $U \subset R^m$ is a parallelopiped then there exists a C^1 function $\varphi : R^m \to R^m$ such that $\varphi([-1, 1]^m) = U$. In addition ψ is one-to-one.

<u>Proof</u>: Denote the $m+1$ vectors which define U by $w^0, w^1, \ldots, w^m$. Define $\varphi : R^m \to R^m$ as

$$\varphi(z) = w^0 + \frac{1}{2} \sum_{i=1}^{m} (z_i + 1)w^i$$

where $z = (z_1, \ldots, z_m)$. Certainly φ is C^1 and if $z \in [-1, 1]^m$ then $0 \leq \frac{z_i + 1}{2} \leq 1$ which implies $\varphi([-1, 1]^m) \subset U$. Furthermore, if $u \in U$ then

$$u = w^0 + \sum_{i=1}^{m} \mu_i w^i \quad \text{where} \quad C \le \mu_i \le 1. \quad \text{Define} \quad z = (z_1, \ldots, z_m) \text{ by}$$

$z_i = 2\mu_i - 1$ then $\varphi(z) = u$ which establishes $\varphi([-1, 1]^m) \supset U$.

The fact that φ is one-to-one follows from the linear independence of the set $\{w^1, \ldots, w^m\}$. For, if $\varphi(z^1) = \varphi(z^2)$ then

$$\sum_{i=1}^{m} (z^1_i + 1) w^i = \sum_{i=1}^{m} (z^2_i + 1) w^i$$

which becomes

$$\sum_{i=1}^{m} (z^1_i - z^2_i) w^i = 0$$

and thus $z^1_i = z^2_i$ for $i = 1, 2, \ldots, m$.

Theorem 6.3: If $U \subset R^m$ is a parallelipiped defined by the $m+1$ vectors $w^0, w^1, \ldots, w^m$ in R^m then there exists a C^1 function $\psi : R^m \to R^m$ (i.e., ψ satisfies (5-4)). Such a function is given by

$$\psi(z) = w^0 + \frac{1}{2} \sum_{i=1}^{m} (\sin z_i + 1) w^i$$

where $z = (z, \ldots, z_m)$.

Proof: Let φ denote the function given by lemma 6.3 and, as in the proof of theorem 6.1, define $\gamma : R^m \to R^m$ by

$$\gamma_i(z) = \sin z_i \qquad i = 1, 2, \ldots, m.$$

Certainly $\psi = \varphi \circ \gamma$ is C^1 and since $\gamma(R^m) = \gamma(Z^*) = [-1, 1]^m$ the theorem follows.

For $m \geq 1$ and real number a_i, b_i with $a_i < b_i$ for $i = 1, 2, \ldots, m$ we will designate the set $[a_1, b_1] \times \ldots \times [a_m, b_m]$ as an (m-dimensional) right parallelipiped. Frequently, in the literature on optimal control theory (especially in the literature dealing with linear problems) it is a control region of this type which is considered. This is due both to the mathematical simplicity of such a set and to its physical relevance, i.e., a right parallelipiped corresponds physically to a control system in which the m controllers are free to move independently (of each other) within a range determined by the upper and lower bounds a_i, b_i.

From definition 6.3 it is easy to see that a right parallelipiped is, in fact, a parallelipiped. Specifically, it is the parallelipiped defined by the vectors

$$w^0 = (a_1, \ldots, a_m)$$
$$w^1 = (b_1 - a_1, 0, \ldots, 0, 0)$$
$$\vdots$$
$$w^m = (0, 0, \ldots, 0, b_m - a_m).$$

Therefore in the particular case of a right parallelipiped the function given by lemma 6.3 reduces to

$$\varphi_i(z) = \left(\frac{b_i - a_i}{2} \right) z_i + \left(\frac{b_i + a_i}{2} \right) \qquad i = 1, 2, \ldots, m$$

and as a corollary to theorem 6.3 we obtain the following.

Corollary 6.1: If $U = [a_1, b_1] \times \ldots \times [a_m, b_m]$ then the function $\psi : R^m \to R^m$ given by

$$\psi_i(z) = \left(\frac{b_i - a_i}{2}\right) \sin z_i + \left(\frac{b_i + a_i}{2}\right) \qquad i = 1, 2, \ldots, m$$

satisfies (6-4) with $Z^* = [-\frac{\pi}{2}, \frac{\pi}{2}]^m$ and $p = m$.

As mentioned previously, frequently in optimal control theory one assumes that the control region is a right parallelipiped. Using the function φ defined by lemma 6.3 and letting $\psi = \varphi|_Z$ where $Z = Z^* = [-1, 1]^m$ we find that ψ satisfies (6-2) with $p = m$. This result, in conjunction with theorem 3.4, justifies the frequently made remark that when $U \subset R^m$ is a right parallelipiped one can assume without loss of generality that $U = [-1, 1]^m$. In fact, as lemma 6.3 shows, this remark is true in the more general case where U is any parallelipiped. Furthermore, when P_1 is linear in the controls i.e., when

$$g_i(x, u) = a_i(x) + \sum_{i=1}^{m} b_{ij}(x)u_j \qquad i = 1, 2, \ldots, n$$

then $f_i(x, z) = g_i(x, \psi(z))$ is linear in the controls z. That is, the function φ given by lemma 6.3 transforms a linear problem P_1 with U a parallelipiped into a linear problem P_2 with $Z = Z^* = [-1, 1]^m$.

We remark in passing that there is certainly nothing unique about the function ψ in theorem 6.3. For example, the function defined by

$$w^0 + \frac{1}{2} \sum_{i=1}^{m} \left(1 + z_i \; e^{\frac{1 - z_i^2}{2}}\right) w^i$$

satisfies theorem 6.3 with $Z^* = [-1, 1]^m$. In practice, the form of P_1 may help to determine which ψ function should be used.

In (17), there is a discussion of a derivation of the Bang-Bang principle for a linear, time-optimal system with $U = [-1, 1]^m$. This discussion is in the spirit of chapter 5 and uses the function given by corollary 6.1 and the (implicit) knowledge that the vertices of U correspond to the points $z \in Z$ where $z_i = \pm \frac{\pi}{2}$ for $i = 1, 2, \ldots, m$. This indicates that, at least in principle, the function given by theorem 6.2 could be used to derive the Bang-Bang principle for a linear time-optimal system with U a polyhedron. For this (and other) reasons it is of interest to determine those points of $Z^* = B^r(1, 0)$ which correspond to the vertices of U i.e., to determine those points $z \in Z^*$ such that $\psi(z) = w^i$ for some $i = 0, 1, 2, \ldots, r$. Inspection of the function ψ given by theorem 6.2 reveals that $\psi(0) = w^0$ and that for $z^1 = (1, 0, \ldots, 0), \ldots, z^r = (0, 0, \ldots, 0, 1)$ we have $\psi(z^i) = w^i$, $i = 1, 2, \ldots, r$. In fact, $\psi(z) = w^0$ for all $z \in Z$ with $\|z\| = 0, 2, 4, \ldots$ and $\psi(\bar{z}^i) = w^i$ where $i = 1, 2, \ldots, r$ and $\bar{z}^i = \alpha z^i$ for $\alpha = \pm 1, \pm 3, \pm 5, \ldots$. Thus, ψ maps $B^r(1, 0)$ and successive shells of thickness 2 onto U.

6.4 U is the Continuous Image of a Convex Body

In the paragraph preceding definition 6.1 we mentioned the case where U is the continuous image of a compact convex set. The following result is a consequence of theorems 6.1, 6.2, and 6.3 and

summarizes conditions sufficient to guarantee that there exist an
onto function $\psi: R^p \to U$ which satisfies (6-3) or (6-4).

$\underline{\text{Theorem } 6.4}$: Let $V \subset R^q$ be a convex body and let $\varphi: V \to R^m$
be continuous. Letting $U = \varphi(V)$, there exists an onto function
$\psi: R^q \to U$ satisfying (6-3). If V is a polyhedron with $r+1$ vertices
and if φ is C^1 on an open set containing V then there exists an
onto function $\psi: R^r \to U$ satisfying (6-4). Finally, if V is a
parallelipiped and if φ is C^1 on an open set containing V then
there exists an onto function $\psi: R^q \to U$ satisfying (6-4).

$\underline{\text{Proof}}$: For each of the three sets $V \subset R^q$ considered there exists
an onto function $\psi^*: R^p \to V$ with $p = q$, $p = r$ and $p = q$ satisfying
(6-3), (6-4), and (6-4) respectively, for an appropriate set $Z^* \subset R^p$.
The theorem follows directly using $\psi = \varphi \circ \psi^*$.

7. SUMMARY AND CONCLUSIONS

Based upon several natural equivalence relations defined on a large class of optimization problems an equivalence concept, called Ψ-equivalence, was developed. In terms of this concept the fundamental question became: given an optimization problem, say P_1, under what assumptions is it possible to construct a second optimization problem, say P_2, which is Ψ-equivalent (hence equivalent) to P_1. In very general terms a complete answer to this question was provided by corollary 2.2.

If an optimal control problem (P_1) has $U \subset R^m$ as its (compact) control region and if a second optimal control problem (P_2) has $Z \subset R^p$ as its control region, where P_2 is obtained by transforming P_1 using the onto function $\psi : Z \to U$, then we may apply corollary 2.2. Thus, we obtained that, if ψ maps the set of admissible controls for P_2 (point-by-point) onto the set of admissible controls for P_1, then P_2 is Ψ-equivalent to P_1. From theorems 3.1 and 3.2 it was possible to rephrase this condition as theorem 3.3: namely, if for each bounded measurable function $u : [t_0, t_1] \to U$ there exists a bounded measurable function $z : [t_0, t_1] \to Z$ such that $u = \psi \circ z$ then P_2 is Ψ-equivalent to P_1. Filippov's lemma guarantees the existence of such a function z if there exists a compact set $Z^* \subset Z$ such that $\psi(Z) = \psi(Z^*) = U$ and thus this lemma reduced the question of Ψ-equivalence to a question of the existence of a function ψ satisfying various forms of what was called the basic assumption. These various forms were conveniently summarized at the beginning of Chapter 6.

In chapter 6 we showed that if U is a convex body in R^m then an appropriate ψ exists and in fact P_1 is equivalent to an optimal control problem P_2 in which the control functions are unconstrained (i.e., $Z = R^p$). Furthermore, when U is a polyhedron we may choose ψ to be C^1 and thus P_1 becomes equivalent to a Lagrange problem. As an application of this result we were able to derive the Pontryagin maximum principle using only the standard necessary conditions of the calculus of variations.

While these results constitute an answer to the fundamental question posed earlier, at least in the setting of optimal control and calculus of variation problems with absolutely continuous trajectories, they suggest several further areas of research, a few of which will be mentioned briefly.

If, instead of bounded measurable control functions, one considers controls which are (for example) piecewise continuous then it appears that all of what has been done herein could be redone in this case <u>provided</u> that a "piecewise continuous analogue" of Filippov's lemma is available. This would certainly be of interest since the vast majority of the traditional calculus of variations research and much of the optimal control research has been concerned with the piecewise continuous case.

As indicated by the proof of the maximum principle (chapter 5) there is reason to believe that the method discussed herein will provide a very effective means of relating other known results in the two areas. This should at least be of interest to those people with backgrounds in the

calculus of variations who (by desire or by necessity) are doing research in optimal control theory.

It is a fact that today most of the existing algorithms for solving optimal control problems are based on the explicit or implicit assumption that the control region is open. Therefore, by transforming P_1 with U compact, into P_2 with $Z = R^p$, we may immediately apply the algorithm to P_2, rather than attempting to modify the algorithms to treat P_1 directly. Since it is also a fact that virtually all optimal control problems must be solved numerically, this last research area may prove to be of immediate importance.

8. LIST OF REFERENCES

1. Bliss, G. A. 1935. Calculus of Variations. Open Court Publishing Co., LaSalle, Illinois.

2. Desoer, C. A. 1959. The Bang-Bang Servo Problem Treated by Variational Techniques. Information and Control 2:333-348.

3. Eggleston, H. G. 1958. Convexity. Cambridge University Press, London.

4. Ewing, G. M. 1969. Calculus of Variations with Applications, W. W. Norton Co., Inc., New York.

5. Filippov, A. F. 1962. On Certain Questions in the Theory on Optimal Control. SIAM J. Control vol. 1, no. 1: 76-84.

6. Hermes, H. 1965. The Equivalence and Approximation of Optimal Control Problems. J. Differential Equations, vol. 1, no. 4: 409-426.

7. Hermes, H., and La Salle, J. P. 1969. Functional Analysis and Time Optimal Control. Academic Press, Inc., New York.

8. Hestenes, M. R. 1950. A General Problem in the Calculus of Variations with Applications to Paths of Least Time. Rand Corporation RM-100. ASTIA Document No. AD-112382.

9. Hestenes, M. R. 1966. Calculus of Variations and Optimal Control Theory. John Wiley and Sons, Inc., New York.

10. Hocking, J. G., and Young, G. S. 1961. Topology. Addison-Wesley Publishing Co., Inc., Reading, Massachusetts.

11. Hurewicz, W., and Wallman H. 1948. Dimension Theory. Princeton University Press, Princeton, New Jersey.

12. Jacobs, M. G. 1967. Remarks on Some Recent Extensions of Filippov's Implicit Function Lemma. SIAM J. Control vol. 5, no. 4: 622-627.

13. Lang, S. 1968. Analysis I. Addison-Wesley Publishing Co., Reading, Massachusetts.

14. Letov, A. M. 1960. Analytic Controller Design II. Automation and Remote Control vol. 21, no. 5: 389-393.

15. McShane, E. J. 1940. Necessary Conditions in Generalized Curve Problems of the Calculus of Variations. Duke Math. J. vol. 7: 1-27.

16. McShane, E. J., and Warfield, R. B. 1967. On Filippov's Implicit Function Lemma. Proceeding AMS 18: 41-47.

17. Park, S. K. 1970. Optimal Control Problems as Equivalent Lagrange Problems. NASA Contractor Report, NASA CR-1550.

18. Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V. and Mishchenko, E. F. 1962. The Mathematical Theory of Optimal Processes. Interscience Publishers, New York.

19. Riesz, F., and Sz-Nagy, B. 1955. Functional Analysis. Frederick Unger Publishing Co., New York.

20. Sagan, H. 1969. Introduction to the Calculus of Variations. McGraw-Hill Book Co., New York.

21. Valentine, F. A. The Problem of Lagrange with Differential Inequalities as Added Side Conditions. Contributions to the Calculus of Variations 1933-37, University of Chicago Press, Chicago, Illinois.

22. Young, L. C. 1969. Calculus of Variations and Optimal Control Theory. W. B. Saunders Co., Philadelphia, Pennsylvania.

9. APPENDIX

In chapter 2 we were concerned with <u>global</u> minimization and consequently with the concept of global weak Ψ-equivalence. Throughout this appendix we will be concerned with <u>local</u> minimization and specifically we will attempt to obtain conditions sufficient to guarantee that P_2 is locally weakly Ψ-equivalent to P_1 where P_2 is defined as in the sentence prior to lemma 2.1. As might be expected, these conditions are obtained only by greatly strengthening the properties of $\Psi: \Omega_2 \to \Omega_1$.

Using the notation and terminology of chapter 2, let us, in addition to the assumptions of chapter 2, assume that Y_1 and Y_2 are <u>topological</u> <u>spaces</u> and that T_1 and T_2 are the relative topologies for $\Omega_1 \subset Y_1$ and $\Omega_2 \subset Y_2$. Furthermore, assume that $\Psi: \Omega_2 \to \Omega_1$ is <u>continuous</u>. Let us agree to the terminology that a neighborhood N_α of $\alpha \in \Omega_1$ (N_β of $\beta \in \Omega_2$) is an open set $N_\alpha \in T_1$ ($N_\beta \in T_2$) containing α (β).

<u>Definition 9.1</u>: If for $\alpha^o \in \Omega_1$ ($\beta^o \in \Omega_2$) there exists $N_{\alpha}o \in T_1$ ($N_{\beta}o \in T_2$) such that $F_1(\alpha^o) \leq F_1(\alpha)$ for all $\alpha \in N_\alpha o$ ($F_2(\beta^o) \leq F_2(\beta)$ for all $\beta \in N_\beta o$) then α^o (β^o) is said to be a <u>local optimal element</u>.

We let $\Omega_1^{oo} \subset \Omega_1$ and $\Omega_2^{oo} \subset \Omega_2$ denote the set of all local optimal elements and we observe that $\Omega_1^o \subset \Omega_1^{oo}$, $\Omega_2^o \subset \Omega_2^{oo}$. With a slight abuse of notation we will refer to the problems of finding local optimal elements for Ω_1 and Ω_2 as P_1 and P_2.

<u>Lemma 9.1</u>: If $\alpha \in \Omega_1$ and $\beta \in \Omega_2$ and $\Psi(\beta) = \alpha$ then $\alpha \in \Omega_1^{oo}$ implies $\beta \in \Omega_2^{oo}$.

Proof: By hypothesis there exists $N_\alpha \in T_1$ in which α is locally optimal. Since Ψ is continuous there exists $N_\beta \in T_2$ such that $\Psi(N_\beta) \subset N_\alpha$. Consider $\beta^* \in N_\beta$ and let $\alpha^* = \Psi(\beta^*) \in N_\alpha$. Since $F_1(\alpha^*) \geq F_1(\alpha)$ we have $F_2(\beta^*) \geq F_2(\beta)$ and thus $\beta \in \Omega_2^{OO}$.

Lemma 9.2: $\Psi(\Omega_2) \supset \Omega_1^{OO}$ iff $\Psi(\Omega_2^{OO}) \supset \Omega_1^{OO}$

Proof: The lemma is trivally true if $\Omega_1^{OO} = \emptyset$. Therefore assume $\Omega_1^{OO} \neq \emptyset$ and consider $\alpha^O \in \Omega_1^{OO}$. If $\Psi(\Omega_2) \supset \Omega_1^{OO}$ then by hypothesis there exists $\beta^O \in \Omega_2$ such that $\Psi(\beta^O) = \alpha^O$ and from lemma 9.1, $\beta^O \in \Omega_2^{OO}$ i.e., $\Psi(\Omega_2^{OO}) \supset \Omega_1^{OO}$. The converse is certainly true since $\Omega_2 \supset \Omega_2^{OO}$.

Lemmas 9.1 and 9.2 and their proofs are vertual repetitions of lemmas 2.1 and 2.2 and are true because of the continuity of Ψ. The analogue of lemma 2.3, however, requires more of Ψ than just continuity. With this in mind we define the following terminology.

Definition 9.2: The continuous function $\Psi : \Omega_2 \to \Omega_1$ is locally open (relative to Ω_2^{OO}) if for each $\beta^O \in \Omega_2^{OO}$ there exists a $N_\beta O \in T_2$, in which β^O is optimal, which satisfies $\Psi(N_\beta O) \in T_1$.

Observe that if Ψ is locally open then it is only necessary for some neighborhood of $\beta^O \in \Omega_2^{OO}$, in which β^O is optimal, that its image be open in T_1. Certainly if Ψ is open then Ψ is locally open.

Lemma 9.3: $\Psi(\Omega_2^{OO}) \subset \Omega_1^{OO}$ iff Ψ is locally open.

Proof: If $\Omega_2^{OO} = \emptyset$ the lemma is vacuously true. Consider $\beta^O \in \Omega_2^{OO}$ and suppose $\Psi(\Omega_2^{OO}) \subset \Omega_1^{OO}$. By hypothesis $\alpha^O = \Psi(\beta^O) \in \Omega_1^{OO}$ and thus there exists $N_\alpha O \in T_1$ in which α^O is optimal. Since Ψ is continuous, $\Psi^{-1}(N_\alpha O) \in T_2$ and $\beta^O \in \Psi^{-1}(N_\alpha^O)$. Letting $N_\beta O = \Psi^{-1}(N_\alpha O)$ we have $\Psi(N_\beta O) = N_\alpha O$ which is open in Y_1. Consider $\beta \in N_\beta O$ then $\alpha = \Psi(\beta) \in N_\alpha O$

so that $F_1(\alpha) \geq F_1(\alpha^O)$ which implies $F_2(\beta) \geq F_2(\beta^O)$. Thus β^O is optimal in $N_\beta o$ which implies Ψ is locally open.

Conversely suppose Ψ is locally open and consider $\beta_0 \in \Omega_2^{OO}$. By hypothesis there exists $N_\beta o \in T_2$ such that $\Psi(N_\beta o) \in T_1$ and such that β^O is optimal in $N_\beta o$. Let $\alpha^O = \Psi(\beta^O)$ and define $N_\alpha o = \Psi(N_\beta o)$. Consider $\alpha \in N_\alpha o$ then there exists $\beta \in N_\beta o$ such that $\Psi(\beta) = \alpha$ and since $F_2(\beta) \geq F_2(\beta^O)$ we have $F_1(\alpha) \geq F_2(\alpha^O)$. Thus $\alpha^O \in \Omega_1^{OO}$ and hence $\Psi(\Omega_2^{OO}) \subset \Omega_1^{OO}$.

As the analogue of definition 2.2 we have the following.

Definition 9.3: P_2 is said to be <u>locally</u> <u>weakly</u> <u>Ψ-equivalent</u> to P_1 iff $\Psi(\Omega_2^{OO}) = \Omega_1^{OO}$.

Theorem 9.1: P_2 is locally weakly Ψ-equivalent to P_1 iff Ψ is locally open and $\Psi(\Omega_2) \supset \Omega_1^{OO}$.

Proof: If $\Psi(\Omega_2^{OO}) = \Omega_1^{OO}$ then from lemma 9.3 Ψ is locally open and from lemma 9.2, $\Psi(\Omega_2) \supset \Omega_1^{OO}$. Conversely if Ψ is locally open then from lemma 9.3 $\Psi(\Omega_2^{OO}) \supset \Omega_1^{OO}$ while if $\Psi(\Omega_2) \supset \Omega_1^{OO}$ then lemma 9.2 implies $\Psi(\Omega_2^{OO}) \supset \Omega_1^{OO}$.

While theorem 9.1 gives a necessary and sufficient characterization of local weak Ψ-equivalence, it assumes <u>a priori</u> a knowledge of Ω_2^{OO}. Thus the following corollary is a far more practical result.

Corollary 9.1: If Ψ is (continuous and) open with $\Psi(\Omega_2) = \Omega_1$ then P_2 is locally weakly Ψ-equivalent to P_1.

Proof: If Ψ is open then it is locally open and $\Psi(\Omega_2) = \Omega_1$ implies $\Psi(\Omega_2) \supset \Omega_1^{OO}$ thus theorem 9.1 applies.

Just as corollary 2.2 served as the basis for the study of global Ψ-equivalence in optimal control theory, corollary 9.1 would play the same

role in a study of local Ψ-equivalence. Certainly, however, the construction of a function Ψ mapping the space of admissible control function Ω_2 into the space of admissible control functions Ω_1 which is at the same time <u>onto</u>, <u>continuous</u> and <u>open</u> appears to be, at best, a difficult job.

In closing let us apply the results of chapter 2 and this appendix to the problem of minimizing a function of n-variables. Specifically, let $Y_1 = R^n$ and $Y_2 = R^p$. From corollary 2.2 we obtain the following.

<u>Corollary 9.2</u>: Corresponding to $\emptyset \neq \Omega_1 \subset R^n$ and $F_1 : \Omega_1 \to R$ suppose there exists a positive integer p, a nonempty set $\Omega_2 \subset R^p$ and a function $\Psi : \Omega_2 \to \Omega_1$ with $\Psi(\Omega_2) = \Omega_1$. In this case $\Psi(\Omega_2^o) = \Omega_1^o$ where Ω_1^o, Ω_2^o denote the set of <u>global</u> minimal elements of $F_1 : \Omega_1 \to R$ and $F_2 = F_1 \circ \Psi : \Omega_2 \to R$.

Notice that Ψ need not be continuous. An obvious application of corollary 9.2 is to construct $\Omega_2 = R^p$ and thus to replace the problem of minimizing $F_1(x)$ subject to the constraint $x \in \Omega_1$ by the problem of minimizing F_2 subject to no constraints. If one is concerned with local minima then corollary 9.1 applies and one obtains the following.

<u>Corollary 9.3</u>: In addition to the hypotheses of corollary 9.2 if Ψ is open and continuous then $\Psi(\Omega_2^{oo}) = \Omega_1^{oo}$ where Ω_1^{oo}, Ω_2^{oo} denote the set of <u>local</u> minimal elements.